PROCEEDINGS of The Institute of Kadio Engineers



Institute of Radio Engineers Forthcoming Meetings

ROCHESTER FALL MEETING Rochester, N. Y. November 14, 15, and 16, 1938

> CINCINNATI SECTION November 11, 1938

CLEVELAND SECTION November 24, 1938

DETROIT SECTION November 18, 1938

INDIANAPOLIS SECTION November 18, 1938

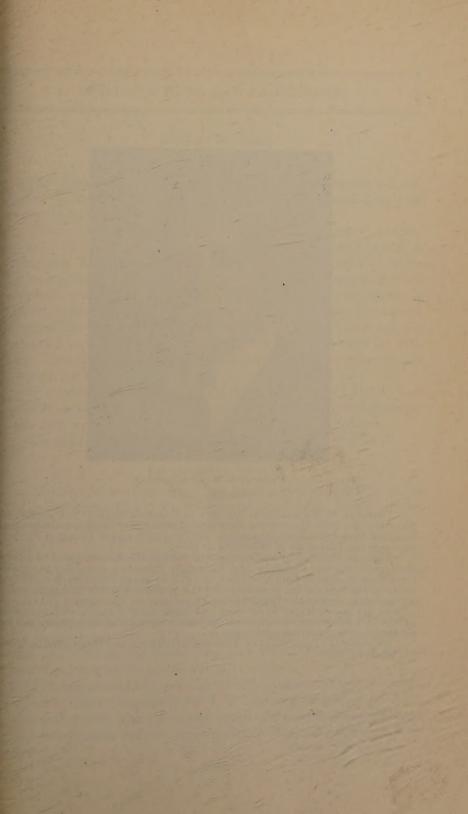
LOS ANGELES SECTION November 15, 1938

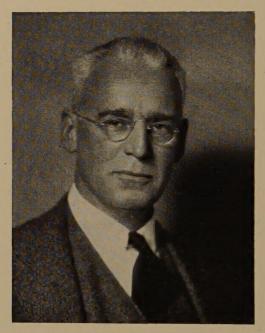
NEW YORK MEETING December 7, 1938

PHILADELPHIA SECTION
December 1, 1938

PITTSBURGH SECTION November 15, 1938

WASHINGTON SECTION November 14, 1938





FREDERICK W. CUNNINGHAM Director, 1938

Frederick W. Cunningham was born in Hamilton Square, New Jersey, on November 22, 1885. He received the degree of Bachelor of Science and Electrical Engineer from Princeton University in 1907 and 1909 and of Doctor of Philoso-

phy from the University of Wisconsin in 1911.

This was followed by engineering activities in the Edison Lamp Works of the General Electric Company, the Winchester Repeating Arms Company, and the Edison Storage Battery Company. In 1922, he joined the radio-development group of the engineering department of the Western Electric Company which subsequently became the Bell Telephone Laboratories. Throughout this connection has been engaged in advisory and supervisory contacts with broadcast tion he has been engaged in advisory and supervisory contacts with broadcast groups and police radio-communication officers as installation engineer and assistant broadcast-development engineer.

He joined the Institute as an Associate in 1926 and was transferred to Mem-

ber grade in 1928.

INSTITUTE NEWS AND RADIO NOTES

Institute Meetings

CHICAGO SECTION

The Chicago Section met on September 9 in Fred Harvey's Union Station Restaurant. V. J. Andrew, vice chairman, presided and 150 were present.

A paper on "Broad-Band Radio-Frequency Transformers" was presented by S. G. Lutz, professor in the School of Engineering of Purdue University. The paper covered a convenient method of calculating the frequency-response, phase shift, and impedance characteristics of tuned coupled-circuit transformers from generalized network equations. It included effects of detuning, dissymmetrical terminating resistances, interwinding capacitance and conductance, mutual resistance, and of a resistance-terminated tertiary winding. The method is valid for both audio-frequency and intermediate-frequency transformers having any coupling coefficients.

It is most useful for coupling coefficients between 0.1 and 0.95, as simplifying assumptions usually made in conventional calculations are valid only for very tightly or very loosely coupled transformers.

The paper was discussed by Messrs. Johnson and Polydoroff.

Los Angeles Section

The May 18 meeting of the Los Angeles Section was held at the Columbia Square Building in Hollywood. Fred Albin, vice chairman,

presided and there were 380 present.

A paper on "Audio-Frequency Facilities of the Western Division of the Columbia Broadcasting System" was given by L. H. Bowman, Columbia western division engineer. The audio-frequency equipment of each of the seven studios and the one large auditorium was described. Salient points of design for flexible and efficient operation were brought out. A new type of volume indicator was stressed and its advantages over former types outlined. Present switching devices and other equipment used in the master control room were discussed in detail. A general discussion followed the presentation of the paper.

"The Acoustical Design of Columbia's Studios" was then discussed by Mr. Spiemyer who spoke for V. O. Knudsen, the designer of the acoustical system. The theoretical design of the studios was discussed and its correlation with the experimentally obtained results was indi-

cated.

Following a general discussion of the paper, an inspection of the facilities was made under the guidance of A. C. Packard, maintenance supervisor.

The June 29 meeting was held jointly with the Society of Motion Picture Engineers at the Filmart Theater in Hollywood. There were thirty-five present.

"Transmission of Motion Pictures over a Coaxial Cable" was the subject of a paper by H. E. Ives of Bell Telephone Laboratories. It covered the theoretical considerations involved in the transmission of both audio-frequency and video-frequency signals over a coaxial cable. The practical aspects and difficulties were discussed in detail by L. F. Brown, who presented the paper for Dr. Ives. A demonstration of the system was made using a reel of motion-picture film and some portable equipment.

A symposium on acetate recording was given at the July 7 meeting of the section which 140 attended. R. O. Brooke, chairman, presided. The meeting was held in the RCA Manufacturing Company plant in Hollywood.

"The New RCA Recorder and Playback for Acetate Recordings" was presented by Barton Kreuzer of the RCA Manufacturing Company. It covered the problems of acetate recording and the advancements made in the development of new equipment. A demonstration was given in the recording of 78-revolution-per-minute records and immediate playback.

"The New Western Electric Feed-Back Recorder for Direct Recording" was the subject of a paper by Charles Sawyer of Electrical Research Products, Inc. It included a discussion of mechanical and electrical feedback applied to recording equipment. It ended with a demonstration of some new equipment using $33\frac{1}{3}$ -revolution-per-minute recording.

The third paper was on "The Manufacture of Materials for Instantaneous Recording." It was presented by W. L. Vail, a manufacturer's agent. It covered the various materials used for acetate and other recording media.

A general discussion of all of the papers was held.

SAN FRANCISCO SECTION

Carl Penther, vice chairman, presided at the September 7 meeting of the San Francisco Section held in Manning's Coffee Cafe. There were thirty-one present.

This meeting was devoted to the review of two papers which appeared in the August Proceedings. The first of "A High-Efficiency Grid-Modulated Amplifier" was reviewed by J. R. Woodyard, and the other, "High-Efficiency Modulation System" was reviewed by Clark Cahill.

WASHINGTON SECTION

On September 12, at the Potomac Electric Power Company Auditorium, there was held a meeting of the Washington Section which was attended by eighty-five. E. H. Rietzke, chairman, presided.

E. N. Dingley, Jr., of the Bureau of Engineering, Navy Department, presented a paper on "A Unique Instrument Landing System for Aircraft." It employs a series of specially designed horizontal loops lying on or buried near the surface of the ground on the flying field and approach. These loops are excited by a 500-cycle generator and for operation to an altitude of 1000 feet require about two kilowatts. The installation is relatively inexpensive and should be reliable and permanent in operation. The plane equipment would be comparatively light. rugged, and foolproof. A cross-pointer indicator informs the pilot of his position relative to the glide path at all times once the plane enters the range of the equipment. The system has not been checked experimentally but tests and calculations appear favorable.

The paper was discussed by Messrs. Burgess, Dorsey, Guthrie, Rietzke, and others.

Personal Mention

P. A. Anderson has left the RCA Manufacturing Company to engage in private consulting practice at Medford Lakes, N. J.

W. B. Bailey, Lieutenant, U.S.N., has been transferred from the

U.S.S. Hopkins to the U.S.S. Moffett, basing at San Diego, Calif.

Previously with General Household Utilities Corporation, E. C. Carlson is now on the engineering staff of Fairbanks Morse and Company, Indianapolis, Ind.

G. W. Gilman who has been with Bell Telephone Laboratories is now assistant technical representative in Europe of the American Telephone and Telegraph Company, with headquarters in London, England.

F. J. Macedo, previously with the Radio Corporation of New Zealand, is now a radio engineer with Collier and Beale, Ltd., of Well-

ington, New Zealand.

E. H. Pierce, Lieutenant, U.S.N., has been transferred from the U.S.S. *Indianapolis* to the Naval Research Laboratory, Washington, D.C.

H. L. Pitts. Lieutenant Commander, U.S.N., has been transferred from San Diego, California to the New York Navy Yard.

J. W. Sanborn has left the RCA Manufacturing Company at Camden, N. J., to join the RCA Victor Company, Ltd., of Montreal, Que.

F. J. Somers, previously with Farnsworth Television, Inc., is now on the staff of the National Broadcasting Company as a television engineer.

F. E. Spaulding, Jr., formerly with the General Electric Company, has joined the engineering department of the Radiomarine Corporation of America, New York City.

W. G. Wagener has left the RCA Manufacturing Company to join the staff of Heintz and Kaufman, Ltd., at South San Francisco, Calif.

R. B. Woolverton, Captain, U. S. Signal Corps, has been transferred to the Engineering Office at Seattle, Wash.

Errata

Mr. H. S. Loh has brought to the attention to the editors the following corrections to his paper, which appeared on the April, 1938, issue of the Proceedings on pages 469-474.

The title should read

"On Single and Coupled Circuits Having Constant Response Band Characteristics."

On page 470, line 6, the equation should read

$$F = \frac{2\Delta f}{f_0}.$$

Equation (6) should read

$$y_0 = 2s\sqrt{m} \; \frac{nQ_1^2}{1-s}.$$

Equation (9) should read

$$Q_1 = \frac{1}{\sqrt{3}F}.$$

TECHNICAL PAPERS

PRACTICAL APPLICATION OF AN ULTRA-HIGH-FREQUENCY RADIO-RELAY CIRCUIT*

By

J. ERNEST SMITH, FRED H. KROGER, AND R. W. GEORGE (R.C.A. Communications, Inc., New York, N. Y.)

Summary—The utilization of an ultra-high-frequency radio circuit for the transmission of telegraph, teletype printer, and facsimile signals is described. The operating procedure is stressed throughout, particularly with regard to equipment maintenance tests and the methods employed to determine the conditions of the circuit such as degree of modulation, signal-to-noise ratio, etc. Considerations are presented with respect to the most efficient division of the total modulation band into the communication channels, as well as the signal-to-noise ratios required for the different types of service. It is found that fading, static, and weather conditions at these frequencies are of little importance as to their effect on the economic use of the circuit. However, diathermy machines and similar sources of disturbance are troublesome and their effect must be minimized. Experience during the past year and a half indicates that the dependability of the ultra-high-frequency radio circuit is of a high order.

INTRODUCTION

HILE the theory of ultra-high-frequency propagation has been studied for several years, there is little information available on the economic utilization of these frequencies for commercial purposes. To obtain such data, the engineers of R.C.A. Communications, Inc., have designed and installed an experimental three-link two-way ultra-high-frequency radio relay between New York and Philadelphia. For the past year and a half this circuit has been utilized by the traffic department as a regular communication facility. This paper will describe the circuit equipment and the means employed to insure dependable operation.

TRANSMITTER

In general, the transmitter construction followed conventional design. Some features however are unique and worthy of detailed description. A close-up view of the transmitter is shown in Fig. 1. The rack on the left houses the high-voltage direct-current supply. The center unit contains the modulator and five test meters. The tanklike structure at the right houses the resonant line control, master oscillator, and power amplifier.

* Decimal classification: R480. Original manuscript received by the Institute, March 3, 1938. Presented before New York Meeting, April 7, 1937.

The five meters together with their proper cables can be plugged into any of the nineteen jacks visible in the picture. These measurements form a part of the daily transmitter test schedule. Each jack is marked with its normal reading so that incorrect operation is easily determined. The racks are built on rollers and may be pulled forward on special skids thus facilitating test and repair work.

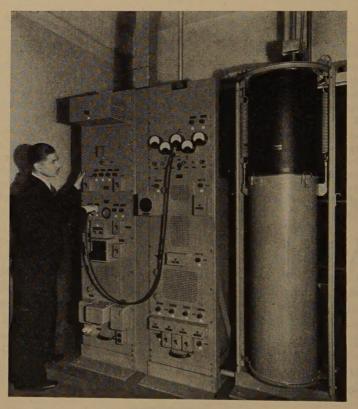


Fig. 1—Front view of an ultra-high-frequency transmitter (installed).

The radio-frequency unit in the tank structure is suspended by springs to reduce vibration. Holes through the black hood allow adjustment of the radio-frequency circuits by means of tuning rods. The small tubular projections fitted with caps and chains permit the insertion of the special thermojunction microammeter shown in Fig. 2. This meter, when held against the glass envelope of a vacuum tube, can be used to predict the life of the tube since experience has shown that subnormal temperature is a sign of low filament emission. The



Fig. 2—Thermojunction microammeter.

upper portion of the radio-frequency unit is painted black since it was found that the difference in heat radiation properties of black and green paint was sufficient to change the original calibration determined for the test model. The locations of the radio-frequency tubes and their associated circuits are illustrated in Fig. 3.

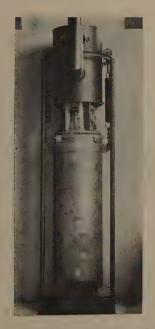


Fig. 3—Radio-frequency unit of an ultra-high-frequency transmitter,

The master oscillator has a tuned grid circuit consisting of a concentric line having a Q in the order of 15,000. A buffer stage was omitted since it was found that without a buffer stage between the power amplifier and the master oscillator there is frequency modulation of only 0.005 per cent at 100 per cent modulation. There is also as frequency shift of the same order when the transmission line is discon-

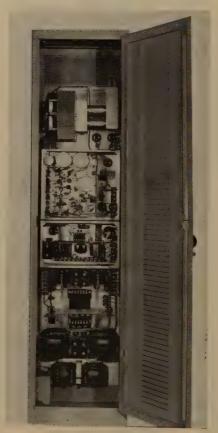


Fig. 4—Rear view of an ultra-high-frequency receiver.

nected or badly matched to the antenna. A set of daily measurements covering a space of two weeks indicated that all transmitters held to a constant frequency within plus or minus 0.005 per cent.

These transmitters are capable of 200 watts output with a carrier frequency of 100 megacycles and a 20-kilocycle modulation band. Normally only half power is used except in extreme conditions of low signal level such as might occur with heavy ice formation on the transmitting antenna.

RECEIVERS

The receivers throughout the system are of the superheterodyne type and utilize loaded-line-section resonators for high-frequency selectivity and oscillator-frequency stabilization. The rear view of a

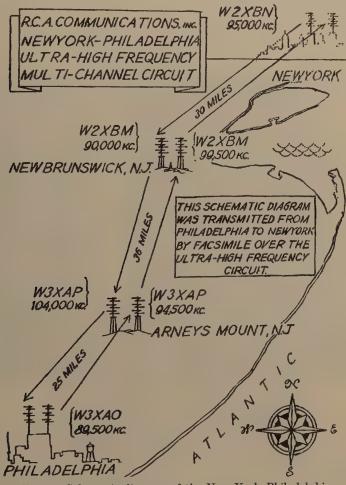


Fig. 5—Schematic diagram of the New York-Philadelphia ultra-high-frequency relay system.

typical receiver unit is shown in Fig. 4. Some features of the receiver

may be of interest.

The high-frequency converter takes the signal from a balanced concentric transmission-line system, through a concentric preselector circuit and a similar concentric resonant circuit to the grid of an RCA 955 tube detector. A stable concentric resonant circuit controls the 955

heterodyne oscillator. The first detector output of 30 megacycles is amplified by two stages using RCA 954 tetrodes, then converted to a lower intermediate frequency of 5.5 megacycles which is amplified before final detection.

Unusually flat automatic gain control is employed to maintain the correct modulation level over the relay circuit. A special resonant circuit and indicator provides means of tuning a signal to the center of the intermediate-frequency amplifier band without interfering with the

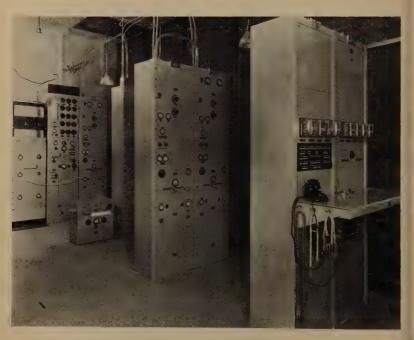


Fig. 6-Interior view of the relay station at New Brunswick.

receiver operation. The operating conditions of the tubes can be determined readily by a group of test meters and switches. All tubes are operated conservatively which materially increases their life.

The power supply is of rugged design and incorporates regulation of the 110-volt alternating-current supply as well as regulation of voltage for the high-frequency tubes and gain-control circuits. The receiver is thoroughly shielded and the construction is such as to permit easy servicing of all parts which are potential sources of trouble. The audiofrequency response of the receiver at present is from 100 cycles to 20 kilocycles with a maximum variation of plus or minus three tenths of a decibel, and with harmonic distortion less than two per cent for the required output level.

OVER-ALL CIRCUIT

A schematic diagram of the over-all circuit is shown in Fig. 5. Two intermediate relay points are used to insure dependable operation. An interior view of one of the relay stations is shown in Fig. 6. The heights of the transmitting and receiving antennas are such that

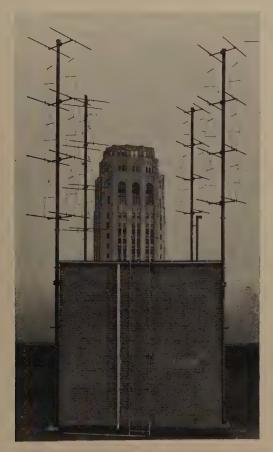


Fig. 7—The transmitting antennas at New York.

the two end relay links are within the optical range and short enough to lay down a strong signal well above the high noise level of the urban terminals. The New Brunswick-Arney's Mount link is just below grazing incidence due to an intervening hill near Arney's Mount. The entire circuit is over land thus permitting the use of horizontal polarization. All antennas are directive; the pine-tree array of horizontal

¹ Bertram Trevor and P. S. Carter, "Notes on propagation of waves below ten meters in length," Proc. I.R.E., vol. 21, pp. 387-426; March, (1933).

dipoles as shown in Fig. 7 is generally employed for both transmitting and receiving although at some points diamond-type receiving antennas are preferred.

The carrier frequencies are staggered according to an allocation plan described in a previous article.2 This arrangement permits a maximum utilization of the available frequency spectrum and minimizes the probability of overlapping or cross talk between the north- and southbound circuits or between links of the same circuit. The Arney's Mount relay is unattended except for weekly inspections and its transmitter frequencies are monitored at New Brunswick to comply with the requirements of the Federal Communications Commission.

This radio circuit has been in operation for approximately 5000 hours. During that time several circuit failures have occurred. The total percentage outage time of the circuit together with its cause is given below.

Transmitter tubes 0.0067 p				per cent;
Receiver tubes 0.074 p		procedure	0.118	per cent ;
ment	per cent	Power failure	0.0018	per cent :

It is believed that the experience gained during the past eighteen months with this circuit will enable the adoption of an improved daily test schedule which will still further decrease the loss of service time.

DESCRIPTION OF CHANNELS

The division of the 20-kilocycle modulation band into the channels now in use is shown in Fig. 8. The five channels indicated in the upper left section of the figure are standard band-pass filters having an effective band width of approximately 100 cycles.3 The frequencies shown correspond to the tone frequencies employed in audio-frequency carrier-telegraph systems. These values represent the geometric mean frequencies of the upper and lower cutoffs of the respective filters. The phase shift of these filters, and, therefore, of the channels is neither symmetrical nor linear about the mid-band frequencies. The resulting frequency distortion and phase distortion have been studied and the results of these studies will be published in the near future. For these narrow-band channels it has been found more economical to employ two types of filters, one configuration for the receiving type and a simpler configuration for the transmitting type. The lesser discrimination of the transmitting filter is permissible since its main function is

² H. H. Beverage, "The New York-Philadelphia ultra-high-frequency fac-simile relay system," RCA Rev., vol. 1, pp. 15-31; July, (1936). ³ The exact band width may be expressed only for idealized cases. See K. Kupfmueller, "On relations between frequency characteristics and transients in linear systems," Elec. Nach. Tech., January, (1928).

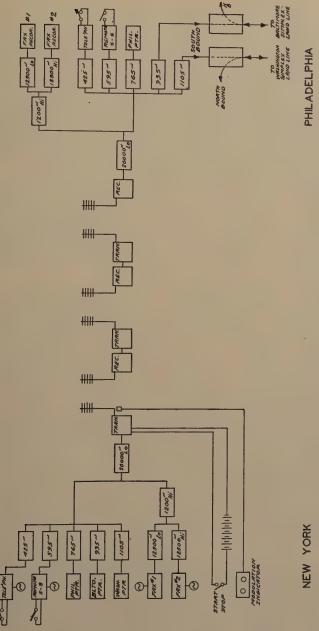


Fig. 8—Schematic diagram of terminal equipment for one direction. Duplicate equipment is used for opposite direction.

to prevent the appearance in the receiving channels of beat frequencies produced between modulation components of the several keyed tones.

The remainder of the modulation band from 1200 cycles to 20 kilos cycles is more or less flexible for facsimile transmission. Possible divisions of this spectrum will be discussed in subsequent sections. For the present, this range has been divided into two wide-band channels. One channel is determined by a 1200-cycle high-pass filter and a 12,500-cycle low-pass filter. The other channel is determined by a 12,500-cycle high-pass filter and a 20,000-cycle low-pass filter. These filters have a discrimination on the order of 60 decibels for frequencies in their attenuating ranges. The 20-kilocycle low-pass "roofing" filter is produided to insure a definite upper limit to the modulation band.

In order to prevent undesired frequency distortion, the modulation band from 300 cycles to 20 kilocycles was designed to have a flat attenuation-versus-frequency response plus or minus one decibel for the over-all circuit. The over-all characteristic was compensated by a two-section lattice-type equalizer installed at the incoming terminals of the north- and southbound circuits.

CHANNELING CONSIDERATIONS

In multichannel communication systems of this type, the division of the modulation band is of primary importance. Where both narrows and wide-band channels are to be provided, it is more economical from the standpoint of filter design to place the narrow-band channels at the low-frequency edge of the available band. This is due to the fact that for a particular filter configuration, the steepness of cutoff, i.e., the discrimination between attenuating and pass-bands, varies inversely as the cutoff frequency. Simpler and, consequently, less expensive band-loss filter configurations can be employed at low frequencies for the same discrimination. This arrangement has the further advantage on less frequency wastage between channels.

Wherever phase or frequency compensation is required for a particular channel, it is generally more economical to design compensating networks for that channel than to redesign the channel filter to have improved characteristics. These improved characteristics require special filter elements such as very high-Q inductances. High-Q elements are more susceptible to mechanical shock and to temperature variations. It is difficult to design inductances with alloy cores having modulation products below minus fifty decibels. High-quality channels of very low noise level should therefore utilize air-core inductances and a more efficient division of the band may be made with these channels at the higher frequencies where higher values of Q are automatically obtained.

TONE-LEVEL-SETTING PROCEDURE

In a central control terminal where both land-line and radio facilities transmit the outgoing signals, it is desirable that the characteristic impedance and required tone levels be identical at the master control board. This arrangement makes for flexibility since no manual adjustments are required by the operator. Moreover, a simple procedure should be sufficient to check a particular type of terminal equipment. Also, in a multichannel setup, the design should be such that maximum energy will be obtained in an individual channel for a fixed peak value of combined tone amplitude. For the ultra-high-frequency radio circuit this peak value should correspond to 100 per cent modulation.

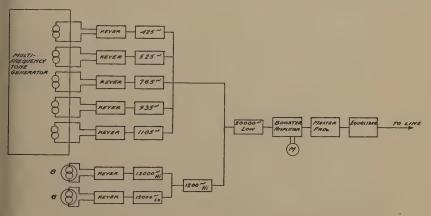


Fig. 9—Schematic diagram to illustrate tone-level procedure.

Obviously, the combined peak value is a function of (a) the number of tones, (b) their phase relationships, and (c) the duration of mark on each channel.

The terminal equipment for the ultra-high-frequency control is shown in Fig. 9. All of the keyer units, whether they be for telegraph, Teletype printer, or facsimile transmission, are adjusted to the same output level. This level is standard throughout the central office so that any service may be transmitted over any facility. In order that this tone level will have its maximum permissible amplitude at the line (or transmitter) input, the booster-amplifier—master-pad combination is adjusted for the number of channels provided by the filter group. When less than the total number of channels are used the master pad may be calibrated against the number of channels. For routine checking of the terminal equipment the control operator applies a single steady tone from one keyer unit. The booster-amplifier output as indicated on a rectox-type meter should then read the value specified for

that terminal. By inserting each keyer unit in turn he can check the complete terminal in a few seconds' time.

The calibration of the master pad depends upon the maximum peak amplitude of the combined tone output which in turn depends upor the three conditions stated above. Neglecting condition (c) for the moment, if n tones of equal amplitudes and random-phase relationships were employed, i.e., n nonsynchronized tone sources, they would all add in phase at some instant thus producing an instantaneous peak amplitude of n times a single-tone value. The master pad should ther have a loss corresponding to a voltage reduction of 1/n.

In practice, the in-phase peaks of short duration may be threshold limited to a certain value depending upon the amount of distortion permissible in the various channels.

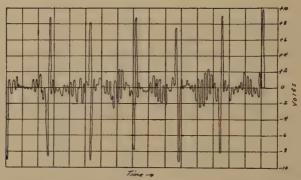


Fig. 10—Resultant of 12 combined tones of a multitone generator.

If, however, the phase relationships of the various tones are fixed as would be obtained for synchronized oscillators or multitone generators the conditions are quite different. In order to present this analysis clearly, a practical example is worked out in detail.

Suppose a multitone generator produces twelve frequencies 450-630, 810, etc., when rotating at 1800 revolutions per minute or 30 revolutions per second. We observe immediately that the armature windings are in the ratio 15, 21, 27, etc., so that we need to calculate only the vector sum of the 12 tones for one revolution of the armature Fig. 10 shows the value of the twelve combined tones during one revolution on the basis that the tones are arranged to produce zero amplitude when all are in phase. The ordinates are in multiples of a singletone amplitude. We see that the master pad would need to reduce the combined tones 1/9.5 or 19.6 decibels, whereas for random tone sources the reduction would be 1/12 or 21.6 decibels. This means the multitone generator could furnish a 2-decibel greater level in each channel than would be permissible with random-phase oscillators.

The above analysis was made on the basis that continuous tones were supplied to all channels. Actually, the channels will be keyed on and off according to the type of communication transmitted. It is apparent that the number of combined peak maxima occurring during an interval of time will decrease as the continuous tones are replaced with keyed tones.

To obtain an idea of the effect of keying on the peak-signal amplitude, we shall continue the analysis of the 12 tones considered above. Let us assume these 12 tones are to be used for 12 telegraph channels and that they are all keyed at 50 words per minute or its equivalent, 20 cycles per second. An analysis of a typical message indicates the signal "on" time to be very nearly equal to the signal "off" time for the international Morse code. The probability that channel No. 1 will be "marking" at a particular instant is then equal to $\frac{1}{2}$. The probability that this channel will be marking over a finite interval Δt is somewhat less than $\frac{1}{2}$, but for our purpose Δt will always be of small magnitude compared to the marking interval.

We may then write

$$P \gg \frac{1}{2}$$

where P is the probability that one channel will be on mark over the interval Δt . The probability that two channels will be marking over the interval Δt is $(1/2)^2$. For n channels the probability that they will all be marking over an interval Δt is

$$P_n \gg (\frac{1}{2})^n$$

or, in other words, one keying maximum would occur every 2^n cycles of the keying frequency. If, for steady tones, one tone maximum of duration Δt occurs in the combined tones every t seconds, the probability that the two maxima will occur simultaneously for n keyed tones is

$$P_k \gg \frac{\Delta t}{t2^n}$$

or one maximum occurs for every $2^n t/\Delta t$ cycles of the keying frequency.

We may then write

$$T = \frac{2^n t t_1 \text{ hours}}{3600 \Delta t}$$

where,

T = time required for 1 maximum

n=number of telegraph channels

t=time for 1 maximum with n steady tones

 t_1 = time of 1 keying cycle

 $\Delta t = \text{duration of peak maximum.}$

For our example above, all channels are keyed at 20 cycles per second and the keying relationships between channels are random. We have n=12 and $t_1=0.05$ second. From Fig. 10, t
ightharpoonup 12.5
extra term to the second are the second at the second are the second at the s

Substituting

$$T = \frac{2^{12}(12.5)(0.05)}{3600} = 1.06 \text{ hours}$$

we observe that for the example given when all 12 tones are keyed a peak maximum occurs on the average of once every hour. It is clean that partial elimination of these occasional peaks by limiting would produce very little distortion in the received channels, and, at the same time, would allow a greater energy level to be transmitted in the individual channels. The lowest limiting threshold value that can be permitted depends upon the rate of recurrence of the peak maxima, the duration of the peak impulse, the frequency spacing of the channels, and the distortion tolerance of the receiving equipment. It is therefore best determined by experiment. Moreover, in practice, long marking intervals will occur on some channels so that the actual peak amplitude will occur more frequently than is indicated by the idealized example above. It is clear, however, that higher audio-frequency signal levels on a greater number of useful channels can be made available for a givent transmitter power by use of the limiting principle.

SIGNAL-TO-NOISE MEASUREMENTS

An important figure of merit of a transmission channel is its signal-to-noise ratio. A particular type of recording device has a minimum signal-to-noise value below which failures occur. In designing a channel for a certain facility it is necessary to compromise between the band width desired for fidelity of the received signal and that required to maintain a usable signal-to-noise ratio. Recent studies indicate that impulse noise varies directly with the band width and fluctuation or smooth noise varies directly with the square root of the band width. Preliminary tests showed the audible noise over the ultra-high-frequency circuit to be generally of the fluctuation type except during occasional intervals of diathermy or motor-ignition interference, at which time the impulse-type of noise predominated.

In order to obtain quantitative noise data over the radio circuit a set of measurements was made. Tones were applied at Philadelphia in various combinations and the outputs of the different channels at New York were compared for signal levels, cross talk, and noise values. Some of these results are shown in Tables I and II.

TABLE I SIGNAL-TO-NOISE MEASUREMENT

Channel	Channel Output at New York in Decibels			
. (cycles, narrow band)	(a) No Signal at Philadelphia	(b) Single 595-cycle tone (100 per cent modulation)		
425 595 765 935 1105 1275 1445 (kilooycles, wide band) 3–17	59 58 60 56 58 59 59	-43 0 -42 -58 -54 -52 -58 -35		

TABLE II

OBSERVATION FOR EFFECT OF CROSS TALK INTO 595-CYCLE CHANNEL FROM OTHER CHANNELS

6, .	Output of 95-cycle channel Decibels
425-cycle tone, 14 decibels below 100 % modulation, steady mark	42
595-cycle tone, 14 decibels below 100% modulation, steady mark	0
935-cycle tone, 14 decibels below 100% modulation, steady mark	42
8-kilocycle tone, 14 decibels below 100% modulation, steady mark	43
425-cycle +935-cycle tone, 14 decibels below 100 % modulation, steady mark	42
425-cycle +8-kilocycle tone, 14 decibels below 100 % modulation, steady mark	42
935-cycle +8-kilocycle tone, 14 decibels below 100% modulation, steady mark	43
425-cycle tone keyed by hand	41
935-cycle tone keyed with RY's on printer	42
8-kilocycle tone keyed by facsimile	42
425-cycle +935-cycle tone keyed by hand and printer	40
425-cycle +8-kilocycle tone keyed by hand and facsimile	
935-cycle +8-kilocycle tone keyed by printer and facsimile	42

It was desired to use one channel of the system as a daily indicator of the noise condition on the circuit. Table I, the second column gives the noise levels in the narrow- and wide-band channels when no signal is applied at Philadelphia. Table I, the third column shows the corresponding measurements with a 595-cycle tone modulating the Philadelphia transmitter 100 per cent. It is seen that the discrimination of the adjacent-channel filters is not sufficient to reduce the 595-cycle signal level below their normal noise levels. This cross-talk value would be still further increased if the 595-cycle tone were keyed. It is therefore apparent that none of the narrow-band channels can serve as a criterion of the circuit noise. The wide-band channel on the other hand has an inherent noise level greater than the cross talk from the 595-cycle tone. Subsequent measurements indicated the wide-band channel to be satisfactory for this purpose.

Table II shows the cross talk obtained in the 595-cycle narrow-band channel for various keyed and steady tones. For these filters it is seen that keying produces about one decibel greater cross talk in the adjacent channels than steady tones. The other channels are generally unaffected due to their greater "frequency distance."

As stated above, the general noise condition of the ultra-high-

frequency circuit is believed to be determined by smooth-type noise. To check this we may use the relation that the noise varies as the square root of the band width. In Table I the wide-band channel is 14 kilocycles and the narrow-band channels are approximately 100 cycles. We should accordingly expect a noise increase in the wide-band channel of 11.83:1(21.5 decibels) which agrees closely with the measured values.

During the occasional intervals when impulse noise predominates, which occurs when diathermy or motor ignition is picked up at any of the transmitters, the wide-band channel may be rendered useless for facsimile communication. The interference appears as wavy diagonal lines in the received facsimile copy. This type of interference does not disturb the normal operation of the narrow-band channels.

Normal atmospherics are of little importance at these frequencies. The automatic volume control incorporated in the receivers compensates for all signal variations due to weather conditions except for exceedingly heavy ice formation on the antennas.

The effects of fading may be considered negligible. During a year's operation of the circuit, three short fading periods occurred, each of about two minutes duration. They are known to be in the New Brunswick-Arney's Mount radio link, and, apparently, have no correlation with the fading periods of normal short-wave frequencies. One theory explains this phenomenon as being due to cancellation of the refraction field by the diffraction field. However, it is not felt that sufficient data are yet available to furnish a complete explanation.

On the whole, the ultra-high-frequency radio circuit has been found quite satisfactory for communication purposes and it is expected that wider applications will demand an increasing utilization of this frequency range.

THE PROBLEM OF SYNCHRONIZATION IN CATHODE-RAY TELEVISION*

By

F. J. BINGLEY

(Philco Radio and Television Corporation, Philadelphia, Pennsylvania)

Summary—This paper considers the effect of various types of transmitted synchronizing signals on the synchronization of television receiver circuits, with particular reference to a system operating on the present standard of 441 lines, 60 fields per second, interlaced scanning.

Starting with a description of interlaced scanning, the various causes which can be responsible for pairing are described. A description of the nature of the transmitted signal and its method of utilization at a television receiver is followed by a discussion of the horizontal synchronizing signal wave shape.

The effect of deflecting circuits on interlacing and the requirements to be fulfilled by the vertical synchronizing signal to ensure interlacing are examined. The various types of vertical synchronizing signals are discussed and compared.

The conclusion is reached that the "narrow vertical" signal is best adapted to ensure good synchronization, good interlacing, and simple, cheap receiver circuits.

Introduction

SINCE a television system is no better than its synchronizing, we must be sure that we recommend the best synchronizing system available when we formulate television standards for the United States. The use of 441-line, 60-field interlaced pictures imposes severe requirements on whatever synchronizing system is selected. The type of synchronizing signal which is adopted will govern the quality of interlacing on the received picture. This, of course, is a very important factor in determining image detail in a high-fidelity system. One glance at the television picture will ordinarily tell whether or not it is synchronized but closer scrutiny is necessary to observe the accuracy of interlacing. Inaccuracy here will result in loss of picture detail. Therefore, we may say that failure to interlace is evidence of poor synchronization.

Appreciating that the solution of this important problem forms the foundation for a successful television system, research work has been conducted at Philco for the past several years on several different synchronizing systems. It is the purpose of this paper to report on this work and to indicate the system which we consider best.

1327

^{*} Decimal classification: R583. Original manuscript received by the Institute, February 17, 1938. Presented before Rochester Fall Meeting, November 8, 1937.

DESCRIPTION OF INTERLACED SCANNING

The method by which interlaced scanning is produced is illustrated in Fig. 1. A small number of lines is chosen so as to simplify the illustration; the results deduced will apply to any odd-line interlaced system. For simplicity, the time occupied in return of the scanning beam both from right to left, and from bottom to top, is regarded as negligible. The directions of scanning shown are chosen in accordance with Radio Manufacturers' Association Standards. The scanning oper-

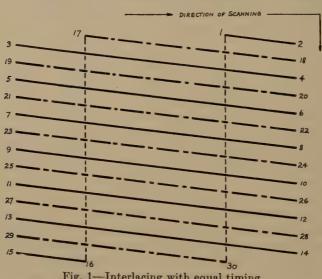


Fig. 1-Interlacing with equal timing.

ation starts at some arbitrary point 1 near the top right-hand corner of the picture, and proceeds with both horizontal and vertical components of motion to point 2. At this time the beam, in response to a horizontal synchronizing impulse, retraces rapidly to 3, from which point it proceeds along the path 3-4 parallel to its original path 1-2. The process is repeated until it reaches the point 16 which it does within the time taken to traverse 6½ horizontal scanning lines. At this point the arrival of a vertical synchronizing impulse causes a rapid retrace in a vertical direction to point 17. Now the normal horizontal and vertical scanning velocities are resumed and the beam moves to point 18, the path 17-18 lying midway between 1-2 and 3-4. At 18 it retraces rapidly to 19, midway between 3 and 5, thence proceeding along 19-20 and similar parallel paths until it arrives at 30, which it does 6½ scanning lines after it left 17. Then another vertical impulse arrives and results in returning the beam to the original point 1. The cycle is repeated indefinitely, the number of horizontal lines scanned between points 1 and 30=13 lines. Thus half the total picture lines are scanned in each frame: for convenience one frame is referred to as an "odd" frame, the next as an "even" frame, and so on, alternately "odd" and "even." The corresponding vertical synchronizing impulses are also called "odd" and "even" respectively.

A television system employing interlaced scanning requires extreme accuracy in the relative placing of successive scans, and the greater

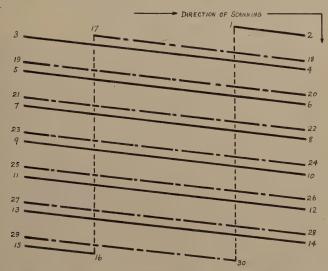


Fig. 2—Pairing with equal timing, due to unequal amplitudes.

the number of scanning lines, the greater is the accuracy required; thus, the demands of a modern 441-line system are extremely rigorous. For example, if we suppose a received image of size $7\frac{1}{2}\times10$ inches, the placing of scanning lines must be accurate to about 0.002 inch for good interlacing. This corresponds to less than 0.03 per cent. When the scanning lines of one frame do not lie accurately in the middle of the gaps left by the preceding one, picture detail is lost and the scanning lines are said to be paired.

CAUSE OF PAIRING

Let us consider Fig. 2. This has been drawn to have the same vertical timing as Fig. 1, namely, $6\frac{1}{2}$ lines between successive verticals, but the amplitudes of vertical deflection for odd and even frames are now no longer equal, as they were in Fig. 1. That is to say, in Fig. 2 the distance 16-17 is not equal to the distance 30-1. This shows that, even

though the vertical timing intervals be equal, interlacing is not achieved if successive amplitudes of vertical deflection are not equal.

Thus we see that the primary cause of pairing is the *inequality of amplitudes of successive* vertical deflections.

THE SYNCHRONIZATION OF RECEIVER CIRCUITS

Let us now examine the manner in which the synchronizing signals are transmitted to the receiver, and the process by which they are utilized at the receiver.

1. Transmission of Synchronizing Signals

After the video-frequency signal leaves the camera it is necessary to modify it before it is suitable for transmission. In the first place, it is desirable to insert a black reference level, which is provided by blanking signals occurring at the end of each line and frame. In addition to this, synchronizing signals are inserted during the blanking periods.

Blanking signals are useful because ordinary types of amplifiers do not transmit absolute voltages, but merely relative values. Thus some reference voltage to indicate black in the picture is essential if the brightness of the receiver screen is to follow faithfully that of the scene being transmitted. In addition to serving as this black reference, these blanking signals may extinguish the scanning beam at the receiver during its retrace between consecutive scanning lines and frames.

Synchronizing signals should be inserted as near the beginning of the blanking signal as possible, so as to leave as much time as possible available for the retrace. For convenience, we shall call the blanking signal occurring at the end of each scanning line the "horizontal blanking impulse," and that occurring at the end of each frame the "vertical blanking impulse." The corresponding synchronizing signals will be referred to as "horizontal synchronizing signal" and "vertical synchronizing signal" respectively.

An oscillogram showing the wave form of the signal corresponding to several lines is shown in Fig. 3. Here negative voltages correspond to white portions of the picture. This is known as a negative videofrequency signal. The synchronizing signals are always inserted in the direction opposite to white, sometimes called the "blacker-than-black" direction, so that they will leave no visible trace on the receiver screen. Thus on the oscillogram shown, synchronizing impulses are in the positive direction, picture signals in the negative direction.

The horizontal blanking impulses terminate at the level marked "black," while the narrower impulses shown above them are horizontal

synchronizing signals. Their usual amplitude is one quarter of that corresponding to the voltage interval between "black" and "white" as illustrated in Fig. 3, this amount having been found in practice to provide enough synchronizing signal to give reliability.

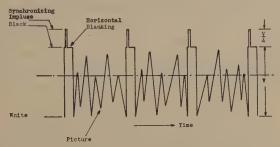


Fig. 3—Composite wide-frequency signal.

2. Separation of Horizontal and Vertical Synchronizing Signals at the Receiver

What are the requirements which synchronizing signals must fulfill in order to give satisfactory results at the receiver? The signal available at the detector will be of the type illustrated in Fig. 3. In order to provide a synchronizing signal for the receiver deflecting circuits, a simple form of amplitude filter generally affords a thoroughly satisfactory result. This amplitude filter permits all portions of the signal above the level marked "black" to be fed to the syn-

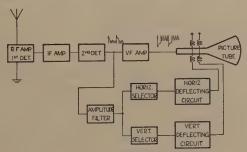


Fig. 4—Television receiver.

chronizing circuits, excluding all trace of the picture signal lying below it. This synchronizing signal contains, of course, both horizontal and vertical signals. It is the duty of the synchronizing circuit to sort out these two sets of signals, and route each to the appropriate deflecting circuit. A diagram of the arrangement of these synchronizing circuits and their relation to the complete receiving system is shown in Fig. 4.

To enable this process of sorting to be carried out efficiently, there horizontal and vertical signals differ in wave form. The wave form of the horizontal and vertical synchronizing signals determines the performance of the synchronizing system in large measure. In particular, the vertical synchronizing signal determines the interlacing quality. In general the output of the horizontal selector has a different wave form from that of the horizontal synchronizing signal; it is convenient to refer to them as horizontal synchronizing impulses. In the same way we shall speak of vertical synchronizing impulses in referring to the output of the vertical selector.

THE HORIZONTAL SYNCHRONIZING SIGNAL

As a result of experience, it has been found that the horizontal signals should not be interrupted even during the transmission of the vertical signal. This must be done to maintain the horizontal oscillator in synchronism at all times. This is true of whatever type of vertical synchronizing that is used: its wave form must be such as not to interfere with the transmission of horizontals; furthermore, it should be of such a shape that simple filters will remove it from the horizontal synchronizing impulses fed to the horizontal deflection circuit. Now let us consider the horizontal synchronizing signal duration and shape. For accuracy it should be steep-sided, but at the same time its frequency spectrum should lie within the band passed by the transmission channel. These considerations indicate for a 441-line system, a duration of 3 microseconds and a time of rise of 1 microsecond.

Considerations of the Effect of Deflecting Circuits on Interlacing

Let us now consider a typical vertical deflecting circuit and examine it critically to determine what factors will influence the interlacing of the deflection pattern for which it, in conjunction with the horizontal deflecting circuit, is responsible.

Such a circuit is shown in Fig. 5: the oscillator, having its frequency control R_1 , is synchronized by a tube T_4 having vertical synchronizing impulses applied to its grid.

The oscillator is coupled to a discharge circuit consisting of a tube T_2 and a plate-circuit network $R_2C_3R_3$ producing a signal which, when fed to the grid of the output tube, causes a saw tooth of current to flow in the vertical deflection coils connected in its plate circuit. The voltage wave form at the grid of T_2 is essentially a sharp positive impulse of duration equal to the time taken for vertical retrace, repeated every 1/60 second.

Examining this circuit, we note that there are just three factors

capable of causing pairing. They are: (1) the effect of inequality of time intervals between successive vertical impulses on the duration or shape or amplitude of the impulse put out by the oscillator T_1 , (2) the effect of inequalities of successive vertical impulses fed to the grid of T_4 on the duration, shape, or amplitude of the impulse put out by the oscillator T_1 , (3) the effect of inequalities of successive vertical impulses on the output of the discharge tube T_2 . Any one of these three factors may be responsible for difference in amplitude of successive vertical deflections due to the corresponding amplitudes of sawtooth deflecting currents fed by tube T_3 into the deflecting coils.

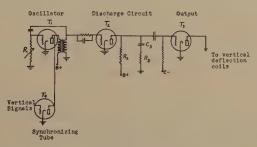


Fig. 5-Vertical deflection circuit.

REQUIREMENTS TO BE SATISFIED BY THE VERTICAL SIGNAL TO INSURE INTERLACING

We are now in a position to state the requirements that the vertical signal adopted must fulfill to insure good interlacing. They are:

(1) Vertical synchronizing impulses for successive frames must be identical in all respects. Note particularly that "impulses" are referred to. The output of the vertical synchronizing separator provides these impulses. If it is of a low-pass filter type, it will have some response, however slight, to the horizontal signals, which will therefore influence the vertical impulse wave shape to some extent. Many expedients have been devised to overcome this difficulty, which will be discussed later.

(2) Precisely equal time intervals between successive vertical

impulses. Again note that "impulses" are referred to.

The fulfillment of these two conditions insures good receiver interlacing. If in addition, the vertical signal is of such a wave shape that the transmission of horizontal synchronizing signals is not interrupted it will be completely satisfactory.

There have been several types of vertical synchronizing signals proposed. How do they compare when examined in the light of the foregoing conditions? Let us consider first what may be called the

simple vertical system.

Types of Vertical Synchronizing Signal

1. Simple Vertical

This was the earliest type of vertical signal. Inserted in a typical television signal, it is illustrated in Fig. 6, which shows both odd and the state of the st

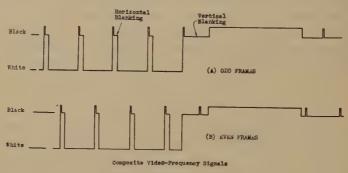


Fig. 6—Simple vertical synchronizing signal.

even frames at (A) and (B) respectively. Note the difference in relative locations of horizontal and vertical signa's in the two frames. This is a necessary condition with an interlaced system of the "odd-line" type. The reason for this difference may be seen by considering a 441-line system as at present in use. In the interlaced system the complete

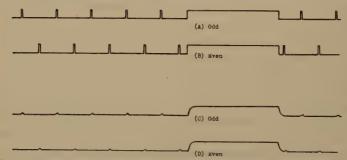


Fig. 7—Resulting vertical synchronizing signal after filtering out the horizontal synchronizing impulses.

picture is scanned in 1/30 of a second, so that in the interval between vertical signals $220\frac{1}{2}$ lines are scanned. Thus the relation between consecutive verticals and the neighboring horizontals differs by $\frac{1}{2}$ line, in odd and even frames.

At the receiver the synchronizing signal may be separated from the picture signal by a device which accepts only that part of the composite picture signal lying above the level marked "black" (Fig. 6) and the resulting synchronizing signal is shown in Fig. 7 at (A)

and (B). This process of amplitude selection is used with all types of synchronizing signal, so that further diagrams of other types of synchronizing signals will show them only after separation from the video-frequency signal has been carried out.

To obtain a vertical synchronizing impulse the combined synchronizing signal shown at (A) and (B) is put through a low-pass filter which gives the result shown at (C) and (D).

It is very difficult to make two such large impulses exactly equal in shape and area, so that as a result this system gives inherently poor interlacing. Furthermore, the long duration of this impulse necessitates the interruption of horizontal synchronizing during the trans-

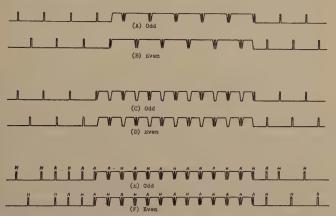


Fig. 8—(A) and (B)—early form of serrated vertical.
(C) and (D)—improved form.
(E) and (F)—serrated vertical with interspersed alternate horizontals.

mission of the vertical signal. We have seen that such omission cannot be tolerated. The system was soon abandoned for these reasons.

Systems capable of providing uninterrupted horizontal synchronizing signals were devised to overcome this difficulty. The two principal types at present used are the serrated vertical system and the narrow vertical system.

2. Serrated Vertical

In the serrated vertical system a vertical signal of the simple vertical type just described is modified by having, as it were, "holes" or "gaps" cut in it to accommodate horizontal synchronizing signals so that they can continue uninterruptedly during the transmission of the vertical.

The earliest form is shown in Fig 8, (A) and (B), where it may be seen that a serious disadvantage lay in the dissimilarity between odd and even verticals, which caused dissimilar outputs from the low-pass filter used to remove horizontal signals. As we have previously seen, pairing resulted.

The signal shown at (C) and (D) was the next step, where twice as many gaps as were necessary to accommodate the horizontal impulses were cut in each vertical, thus rendering them more similar. But the differences in location of horizontal signals in successive verticals caused dissimilarities which resulted in pairing, though not as badly as before.

The final form of this system is shown at (E) and (F). Here extra impulses known as alternate horizontal signals are added during

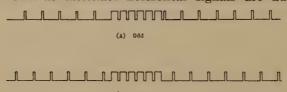


Fig. 9—Form of synchronizing signals used by British Broadcasting Corporation.

the vertical signal and for short periods preceding and following it. The odd and even frames are shown at (E) and (F), respectively, and the locations of horizontal signals and alternate horizontals are indicated by the letters H and A, respectively, placed above them.

In this synchronizing system, horizontal synchronizing impulses are obtained from the combined synchronizing shown in Fig. 8 by means of a high-pass filter which permits only the horizontal impulses to pass.

A modification of the serrated vertical system, which is now being used in the British Broadcasting Corporation's London transmitter, is shown in Fig. 9. It is to be noticed that, during the vertical signal, the horizontal signals have been run over into the serrations at the leading edge. This removes the necessity for alternate horizontals, at least during the vertical synchronizing signal.

3. Narrow Vertical System

The wave form used in this system is shown in Fig. 10. It will be noted that, first, a single impulse only is used; second, it is made narrow enough that it does not interfere with a horizontal signal in either odd or even frames; and, third, the impulse is transmitted at an amplitude 25 per cent greater than the horizontal signals; that is to say, the vertical signal is one and a quarter times the amplitude of the horizontal. This permits amplitude selection to be used for deriving the vertical synchronizing impulse from the combined synchronizing signal.

It can be selected by the use of a simple biased tube as shown in Fig. 10 (C), thus dispensing with low-pass filters and the resulting effecting of neighboring horizontals on the resultant vertical synchronizing impulse. It is the only synchronizing system in which the successive vertical impulses can be made exactly alike.

The horizontal synchronizing impulses are obtained through a high-pass filter, as in the case of the serrated vertical system.

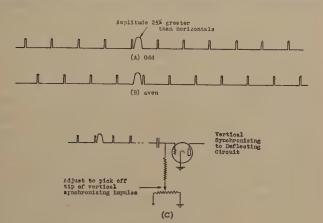


Fig. 10—(A) and (B)—narrow vertical system.
(C)—fundamental circuit for selecting vertical synchronizing impulse.

DISCUSSION OF THE RELATIVE MERITS OF SERRATED AND NARROW VERTICAL SYSTEMS

1. Fulfillment of Requirements for Interlacing

How will these two systems fulfill the two requirements previously proved to be essential for good interlacing? The first condition it will be recalled was that vertical synchronizing impulses for successive frames must be identical in all respects.

With careful adjustment of the amplitude and shape of the alternate horizontal signals, and of the gaps of serrations left in the verticals, this condition can be fulfilled within close limits in the case of serrated vertical. Narrow vertical, of course, fulfills it, and has the added advantage that the synchronizing impulse applied to the vertical deflecting circuits at the receiver has an extremely small area, very much smaller in fact than that of the vertical impulse resulting from the use of a serrated signal.

Now we have seen that differences in shape or area of the vertical impulses will result in pairing through their effect on the oscillator and discharge circuit. Thus a given percentage difference in area between

successive vertical impulses will be much more serious in its effect upon interlacing when the serrated impulse is used.

The second condition was that there must be precisely equal time intervals between successive vertical impulses.

Narrow vertical has the distinct advantage in that it has a sharp wave front, being of such short duration, and thus provides a very accurate time marker. On the other hand, after the serrated signal passes through the low-pass filter used to remove horizontals, it appears as a relatively broad impulse at least six times longer in duration than the narrow vertical signal. Thus the synchronizing is less accurately determined since the wave front is not nearly as steep. Because of this the interlacing is more easily influenced by fluctuations in supply voltages.

Besides the necessary conditions just discussed, there are some further considerations. The narrow vertical system is inherently simpler to produce and control than the serrated vertical system.

The additional amplitude called for by the narrow vertical signal has been cited as a disadvantage. It must be remembered however that this additional amplitude is very small, representing only 6.25 per cent of the total peak-to-peak composite video-frequency signal, and that it does not have to occupy a linear portion of the modulation characteristic.

Consider the case of a 10-kilowatt transmitter where the horizontal and vertical synchronizing signal amplitudes are equal. By how much must the average power be increased if narrow vertical is substituted? The answer depends on the polarity of transmission. With negative transmission, there is no increase in average power required. If the transmission is positive, however, the average power will need to be increased to 11.5 kilowatts.

In present television systems working with alternating-current-coupled video-frequency amplifiers, the background component must be reinserted at the receiver by a device called a leveling tube, which causes the tips of the synchronizing impulses to be held at an approximately constant voltage at all times. Necessarily, some distortion is introduced into the video-frequency signal by this leveling process. This distortion is less when a narrow vertical signal is used.

EXPERIENCES AT PHILCO WITH THE NARROW VERTICAL SYSTEM

Previous to their adoption of the narrow vertical system Philco experimented extensively with the other types of signal described. The narrow vertical has been in use for over two years and has proved

its advantages conclusively during that period. It has been used with both positive and negative transmission with equal success. Its use has enabled the simplification of receiver circuits so that excellent results can be obtained with the minimum of receiver equipment.

This matter of simplification, we all realize, is of vital importance to the future commercial television receiver. We, therefore, believe that the narrow vertical system should be adopted as standard in the United States.

ON THE LONG-PERIOD VARIATIONS IN THE F₂ REGION OF THE IONOSPHERE*

By

KEIKITIRO TANI, YOJI ITO, AND HIROSI SINKAWA (Naval Technical Research Department, Tokyo, Japan)

Summary—The long-period variations in the F2 region are discussed with data based on observations of the ionosphere for the last 42 months since May, 1934, by the Naval Technical Research Department, Tokyo. The variations in the critical frequency of the penetration are presented in the form of monthly averages of the noon values. The mean values of the hourly critical frequencies of the penetration for a month were calculated in order to measure the approximate value of the energy absorbed by the F₂ region. It is shown that these mean values change with the zenith angle of the sun, and they seem to have similar characteristics in both hemispheres. The rate of the seasonal variation in these values, moreover, is much greater in the southern hemisphere than in the northern hemisphere. From this it is presumed that there are two kinds of variations in the energy absorbed by the F₂ region, the one seasonal and the other annual, the effect of these variations being differential in the northern hemisphere and additional in the southern hemisphere. The anomalous variation in the critical frequency of the F2 region may be explained by this presumption with the aid of the hypothesis of thermal expansion.

A graph of the seasonal variations in the minimum virtual height is given, in which it is shown that this variation seems to prove the foregoing presumptions.

The yearly variations in the critical frequencies are examined, and it is shown that there is considerable correlation between them and the sunspot numbers or solar constants. The data, however, are too meager to allow any conclusions to be reached, the period of observation having been only about one third of a sunspot cycle.

ITH the development of the experimental study of the ionosphere, we have now a number of ionospheric regions, namely, the C, D, E1, E2, F1, F2, and G regions. Of these the C, D, E2, and G regions were irregularly observed on special occasions, but except that the D region acts as an absorbing layer, the extent to which they affect actual radio wave transmission is not known.

On the other hand, the characteristics of the E1, F1, and F2 regions, which are stable, were fully investigated, and it was found that the electron density of the E1 and F1 regions varies with the zenith angle of the sun. 1,2,3,4 The daily and seasonal variations in the electron den-

* Decimal classification: R113.61. Original manuscript received by the Institute February 18, 1938. Published in Report of Radio Res. (Japan), vol. 7, pp. 91-96;

October, (1937).

1 J. P. Schafer and W. M. Goodall, "Diurnal and seasonal variations in the ionosphere during the years 1933 and 1934," Proc. I.R.E., vol. 23, pp. 670-700;

June, (1935).

2 S. S. Kirby and E. B. Judson, "Recent studies of the ionosphere," Nat. Bur. Stand. Jour. Res., vol. 14, pp. 469–486; March, (1935).

3 T. R. Gilliland, "Multifrequency ionosphere recording and its significance," Nat. Bur. Stand. Jour. Res., vol. 14, pp. 283–303; March, (1935).

4 T. Minohara and Y. Ito, "Ionosphere studies during the total solar eclipse of June 19, 1936," Report of Radio Res. (Japan), vol. 6, pp. 115–142, (1936).

sity of the F₂ region alone remains unexplained. The hypothesis of thermal expansion of the upper atmosphere, which was invoked to account for these variations in the northern hemisphere, 5,6,7 failed to explain the phenomena that is observed in the southern hemisphere.8 It is very important not only in the study of radio wave propagation, but also for researches in geophysics, to elucidate the nature of the F₂ region. The most effective way in which that could be done is to make observations in widely separated regions throughout the earth's surface. This paper gives the monthly average values of the measurement made since 1934 in Tokyo, and compares them with those of other stations. An attempt was made to determine the long-period variation due to changes in the sun's activities.

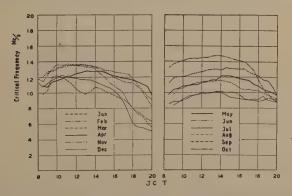


Fig. 1-Monthly averages of the diurnal variation of the critical frequency (xcomponent).

The measurements have been continued at the Naval Research Department (35°38.5' north, 139°42.5' east) since May, 1934, and seeing that the multifrequency measurement by manual method was used. it was possible to determine the critical frequency of penetration and the minimum virtual height of the region. The data are given in monthly averages, thus eliminating the short-time fluctuations. The averages are for one or two days each week. The monthly averages of the diurnal variations in the critical frequency are shown in Fig. 1.

Since the noon values usually represent the maximum values for a day, it was sufficient to use the noon values in the report, except,

⁵ E. O. Hulbert, "Analysis of recent measurements of the ionosphere," Phys. Rev., vol. 46, pp. 822–823; November, (1934).

⁶ E. V. Appleton and R. Naismith, "Some further measurements of upper atmosphere ionization," Proc. Roy. Soc. A., vol. 150, pp. 685–708; July, (1935).

⁷ D. F. Martyn and O. Ö. Pulley, "The temperature and constituents of the upper atmosphere," Proc. Roy. Soc. A., vol. 154, pp. 455–486; April, (1936).

⁸ L. V. Berkner, H. W. Wells, and S. L. Seaton, "Characteristics of the upper region of the ionosphere," Terr. Mag., vol. 41, pp. 173–184; June, (1936).

however, in such a special case as is covered in the discussion which now follows. The seasonal variations in the noon value of the critical frequency are shown in Fig. 2, from which it will be seen that they repeat almost similar seasonal characteristics every year, and that the absolute value increases from year to year. This seasonal characteristic, which is the most outstanding feature of the F_2 region, is almost contrary to those of the E_1 region and the F_1 region, whose characteristics change with the zenith angle of the sun. The hypothesis of thermal expansion, which was formulated by a number of investigators of the data that were obtained in the northern hemisphere. Berkner and others have pointed out that the characteristics in both hemispheres

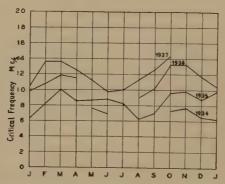


Fig. 2—Seasonal variations of the critical frequency in monthly averages noon value (×component).

are almost alike, so that the relation between the local season and the noon critical frequency becomes reversed. The hypothesis of thermal expansion does not explain these critical-frequency phenomena. Berkner presumed that these phenomena are not seasonal, but that they are an annual effect, which may be attributed to changes in the distance of the earth from the sun, or a change in the mean orientation of the earth's magnetic field, or a change in the position of the earth with respect to solar latitude, or some such change which occurs as the earth moves in its orbit.

Now, in the thermal-expansion theory, it is supposed that the energy absorbed by the F₂ region varies with the zenith angle of the sun, whereas the irregular changes in the critical frequency are due to other causes. In the annual-effect theory, the energy itself is supposed to change irregularly. In discussing the change in energy, it is not sufficient to take the noon average values alone. In this paper the mean values of hourly critical frequencies throughout a month are so ad-

justed that they will denote the approximate relative values of the energy absorbed by the F₂ region during that month. Fig. 3 shows the seasonal variation in this mean critical frequency. This change may be considered to correspond almost to that of the zenith angle of the sun, so that the variation in the energy absorbed by the F₂ region has seasonal characteristics. But it is necessary to examine the data obtained by the same method in the southern hemisphere. According to the results obtained at Watheroo,⁸ and judging from the diurnal variations there in the critical frequencies and their noon values, the characteristics of the mean critical frequencies might be seasonal, as in the northern hemisphere, and the seasonal difference might be greater than in the northern hemisphere. This similarity in seasonal character-

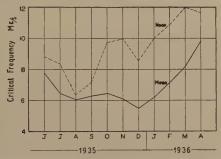
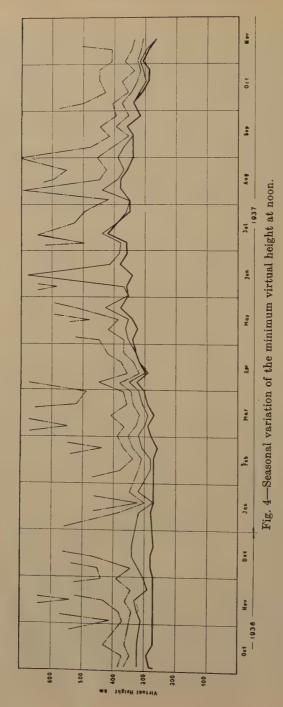


Fig. 3—Seasonal variation of the monthly mean critical frequencies (xcomponent).

istics in both hemispheres and the great difference in the rate of change at two points whose latitudes are nearly equal in absolute values, leads to the following hypothesis. The energy absorbed by the F2 region undergoes two different effects, the one being seasonal, which changes with the zenith angle of the sun, and the other annual, which changes simultaneously in both hemispheres. These effects act differentially in the northern hemisphere, and additionally in the southern hemisphere, causing the seasonal effect to be fairly small in the former and very great in the latter. The seasonal variation in the energy absorbed by the F2 region is so small in the northern hemisphere that the noon value of the critical frequency becomes lower in summer than in winter as a result of the thermal-expansion effect, while in the southern hemisphere the seasonal variation is so great that the noon value becomes higher in summer, outstripping the thermal effect. Thus this presumption, together with the thermal-expansion hypothesis, explains the contrary characteristics of the seasonal variation in the noon value, and the similarity in type of the diurnal variation in the critical frequencies in both hemispheres.



The variation in the "minimum" virtual height, which is the lowest virtual height observed for the region, seems to justify the foregoing presumptions. Fig. 4 shows its variation and some of the virtual heights measured at several frequencies at noon during the last year. It is probably in order to consider the minimum virtual height as representing the true height of the region. The height is lowest in winter, about 270 kilometers, and highest in summer, about 360 kilometers. This characteristic also holds in the southern hemisphere, and it seems to show that the F₂ region has a seasonal character, exactly opposite to the characteristic of the noon value of the critical frequency,

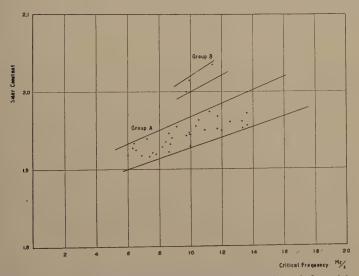


Fig. 5—Correlation of the monthly average noon value of the critical frequency and the solar constant.

although further investigations are needed before it is possible to determine the extent of the correspondence of the minimum virtual height to the true height, or the rate of the effect of the retardation by the F_1 region.

The existence of yearly variations is clearly shown in Figs. 2 and 3. It is reasonable to suppose that this phenomenon is due to the long-period variation in the solar radiation of the type responsible for the ionization of this region. The variation in the monthly average of the noon critical frequency seems to be correlated in some way with the sunspot number as the solar constants. The solar constant as given by Ängström⁹ is

A. Ängström, Astrosphys. Jour., vol. 55, p. 24, (1922).

$S = 1.903 + 0.0055\sqrt{N},$

where S is the solar constant and N is the sunspot number. Fig. 5 shows the relation between the monthly average of the noon values of the critical frequencies and the solar activities. Generally speaking, the critical frequencies seem to change linearly in proportion to the solar constant until the latter becomes very large. It is shown as Group A in the figure. Group B indicates the irregular case when the solar constants are very large. Owing to the yearly irregularity in the seasonal variations the attempt to discover the yearly variation in the critical frequencies for each month of the year was not successful. No conclusion, however, could be reached with the scanty data available, the period of observation being only one third that of the sunspot cycle.

Conclusions

It is concluded that the monthly averages of the noon value of the critical frequency of penetration of the F_2 region undergo variations in two different periods, the one being due to the season and the other to solar activity. The hypothesis that the energy absorbed by the region changes both seasonally and annually, and that they affect differentially in the northern hemisphere and additionally in the southern hemisphere, explains, in conjunction with the thermal-expansion theory, the various characteristics of the F_2 region. The change in the minimum virtual height seems to support the hypothesis. Although the relation between the monthly average of the critical frequency and the solar constants was investigated, no conclusion could be reached, because of too meager data.

──>•≺**○**

MAXIMUM USABLE FREQUENCIES FOR RADIO SKY-WAVE TRANSMISSION, 1933 to 1937*

By

T. R. GILLILAND, S. S. KIRBY, N. SMITH, AND S. E. REYMER (National Bureau of Standards, Washington, D. C.)

Summary—Graphs are presented to indicate the maximum usable frequencies for sky-wave transmission from June, 1933, to December, 1937. The graphs are given for March, June, and December to show equinoctial, summer, and winter conditions for each year. The factors which must be considered in deriving these graphs from vertical-incidence ionosphere measurements are outlined briefly. The principal factors are Snell's law, variation of virtual height of the ionosphere with frequency, the curvature of the ionosphere, and the effect of the earth's magnetic field.

The method of applying these graphs to simple and complex transmission paths

is indicated.

The effect of sporadic E reflections, absorption, and scattered reflections are discussed briefly.

In conclusion it is pointed out that the graphs and the ionosphere data, from which they were derived, may be used to estimate future diurnal, seasonal, and long-time variations of maximum usable frequencies.

I. Introduction

HE graphs of maximum usable frequencies for radio sky-wave transmission presented here were prepared from the vertical-incidence ionosphere measurements made by the National Bureau of Standards at Washington. These graphs, for any particular month, are derived from the average condition of the ionosphere for that month. The graphs give results representative of the winter, summer, and equinoctial periods from June, 1933, to December, 1937, inclusive.

The maximum usable frequency for any distance is defined as the highest frequency which can be used for radio sky-wave transmission over the given distance. Waves of frequencies higher than this penetrate through the ionosphere and are not returned to the earth; waves

pp. 627-640; May, (1938).

1 T. R. Gilliland, S. S. Kirby, N. Smith, and S. E. Reymer, "Characteristics of the ionosphere at Washington, D.C.," Nat. Bur. Stand. Jour. Res., vol. 18, pp. 645-667, June, (1937); Proc. I.R.E., vol. 25, pp. 823-840, July, (1937), and a series of monthly reports beginning in Proc. I.R.E., vol. 25, p. 1174, Septem-

ber, (1937).

^{*} Decimal classification: R113.7. Original manuscript received by the Institute, March 2, 1938. A preliminary version of this paper was presented as paper C6 of the meeting in London, November, 1937, of the Special Committee on Radio Wave Propagation. Publication approved by the Director of the National Bureau of Standards. Published in the Nat. Bur. Stand. Jour. Res., vol. 20, pp. 627-640: May. (1938).

of lower frequencies are reflected and may be used for transmission over this distance. Graphs of maximum usable frequencies in terms of distance also give skip distances as a function of frequency. The maximum usable frequencies give more directly than skip distances information usually desired in connection with radio transmission; i.e., determination of the maximum frequency which can be used over a given distance is of more general value to an engineer than the minimum distance over which communication can be had on a given frequency.

II. RELATION OF VERTICAL TO OBLIQUE-INCIDENCE TRANSMISSION

Radio waves incident obliquely upon the layers of the ionosphere are more easily reflected than those incident vertically, that is, they can be reflected from regions of smaller ionization density. In general, for a given frequency of the transmitted wave, the larger the angle of incidence, the less is the ionization density required for reflection. Thus, if a given layer will reflect waves up to a certain frequency at vertical incidence, it will reflect waves of considerably higher frequencies at oblique incidence.

The angle of incidence is the angle between the ray and the normal to the lower boundary of the ionosphere at the point where the ray enters the ionosphere. The angle of incidence should not be confused with the angle of departure, which is the angle above the horizontal at which the ray is projected from the transmitting station. Although, in general, a transmitting station may radiate energy at all angles above the horizontal, only certain ray paths, corresponding to a few definite angles of departure, are useful for transmission over a given distance at a given time. At vertical incidence the angle of incidence is zero, and increases for oblique incidence.

The angle of incidence depends on (1) the distance of transmission for one reflection from the ionosphere, and (2) the height of the layer. Given these two quantities, therefore, the angle of incidence may be calculated. The maximum usable frequency is given roughly, from this angle and the vertical-incidence critical frequency, by the following relation, known as the "secant law":

$$f'=f_c\sec\phi_1$$

where,

f' = maximum usable frequency

 f_c =critical frequency or highest frequency reflected at vertical incidence from a given layer.

 ϕ_1 = angle of incidence.

The given distance will be within the "skip zone" for all higher frequencies.

The secant law does not take several important factors into account. The first and most important of these is the fact that the virtual height of a given layer of the ionosphere increases with frequency, especially near the critical frequency. As a result it is difficult to choose the proper

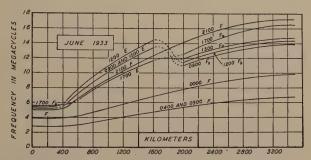


Fig. 1—Maximum usable frequencies, latitude of Washington, D. C., June, 1933.

height to be used in calculating the angle of incidence for a given distance. A method whereby, for a flat ionosphere, this difficulty may be overcome has been described by Smith.²

The second factor which must be considered is the curvature of the ionosphere. This introduces another modification in the calculation of

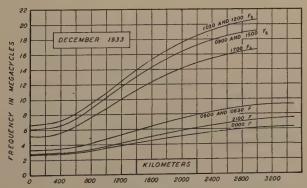


Fig. 2—Maximum usable frequencies, latitude of Washington, D. C., December, 1933.

the angle of incidence for a given distance and a given height of reflection. This results in a higher maximum usable frequency over a given distance than would be calculated on the basis of a flat ionosphere.

Still another factor that must be taken into account is the effect of

² Newbern Smith, "Extension of normal-incidence ionosphere measurements to oblique-incidence radio transmission," Nat. Bur. Stand. Jour. Res., vol. 19, p. 89; July, (1937).

the earth's magnetic field. In the presence of the earth's field a radio wave in the ionosphere is split into two components known as the ordinary ray and the extraordinary ray. These components are reflected from different levels in the ionosphere, and consequently have different critical frequencies. At Washington the vertical-incidence critical fre-

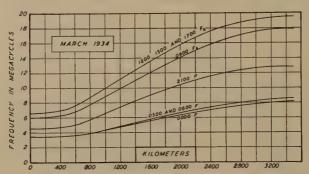


Fig. 3—Maximum usable frequencies, latitude of Washington, D. C., March, 1934.

quency for the extraordinary ray is about 800 kilocycles higher than for the ordinary ray, for frequencies in the neighborhood of 5000 kilocycles and the separation is greater for lower frequencies. At oblique incidence the effect of the earth's magnetic field complicates the

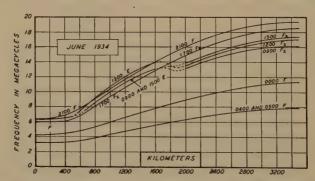


Fig. 4—Maximum usable frequencies, latitude of Washington, D. C., June, 1934,

calculation of maximum usable frequencies. The maximum usable frequency for the extraordinary ray is always greater than for the ordinary ray. Except for short distances, the difference is not great, and decreases with increasing frequency and distance. For practical communication problems it is sufficiently accurate to make the calculations for the ordinary ray.

Methods of obtaining maximum usable frequencies, taking into account the earth's magnetic field and the curvature of the ionosphere have been presented by Smith.3 These methods have been used in preparing Figs. 1 to 14.

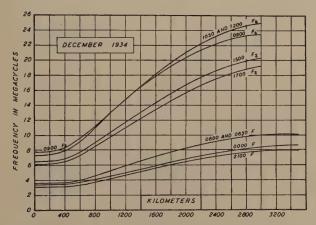


Fig. 5—Maximum usable frequencies, latitude of Washington, D. C., December, 1934.

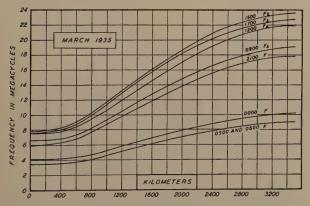


Fig. 6-Maximum usable frequencies, latitude of Washington, D. C., March, 1935.

III. TRANSMISSION PATH

The geographical part of the ionosphere which controls long-distance high-frequency radio transmission is that part traversed by the wave in passing from the transmitter to the receiver. The nearest part

³ Newbern Smith, "The application of vertical-incidence ionosphere measurements to oblique-incidence radio transmission," Nat. Bur. Stand. Jour. Res., vol. 20, pp. 683-706; May, (1938).

of the ionosphere thus traversed may be as much as 1000 to 2000 kilometers from the transmitter, and a similar distance from the receiver. Ionosphere conditions at this point may be considered to be

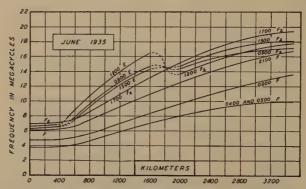


Fig. 7—Maximum usable frequencies, latitude of Washington, D. C., June, 1935.

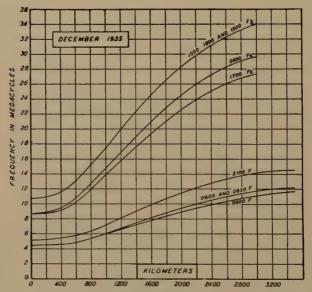


Fig. 8—Maximum usable frequencies, latitude of Washington, D. C., December, 1935.

the same as conditions at the transmitter or receiver, at the same local time, provided no great differences of latitude exist. Consequently, depending on the time of day, different frequencies might be necessary for transmission in different directions.

Take for example, stations at Washington, D. C., North Platte, Nebraska, and San Francisco, California, in the eastern, central, and western United States respectively. The distances between Washing-

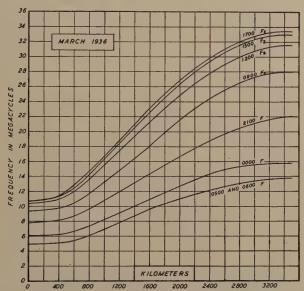


Fig. 9—Maximum usable frequencies, latitude of Washington, D. C., March, 1936.

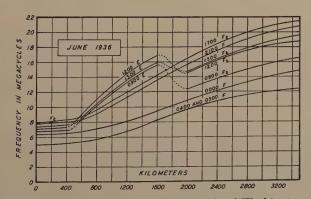


Fig. 10—Maximum usable frequencies, latitude of Washington, D. C., June, 1936.

ton and North Platte and between North Platte and San Francisco are approximately 2000 kilometers. Also take the time as 0715 local time at North Platte in December, 1937 (see Fig. 14). Then the local time at the reflection point between North Platte and Washington

would be about 0800 and the local time at the reflection point between a North Platte and San Francisco would be about 0630. By interpolating linearly between the graphs for successive times in Fig. 14 it is evident that, at this time, the highest frequency (about 19,000 kilocycles) which can be used between North Platte and Washington is much greater than the highest frequency (about 8300 kilocycles) which can be used over the different path between North Platte and San Francisco

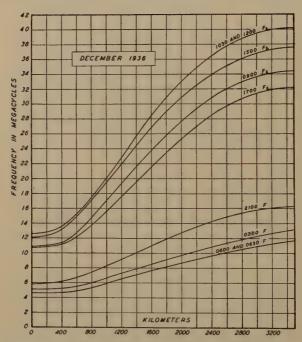


Fig. 11—Maximum usable frequencies, latitude of Washington, D. C., December, 1936.

cisco. This does not mean, however, that different frequencies should be used for two-way communication over the same path such as North Platte to Washington and Washington to North Platte, since the maximum usable frequency is the same for transmission in opposite directions over the same path.

For single-reflection transmission the controlling portion of the ionosphere is halfway between the terminal points. For multireflection transmission the ionosphere along the entire path, except for that portion within several hundred kilometers of the terminal points, must be considered. Because of large differences in local time and latitude encountered in long transmission paths involving more than one reflec-

tion, widely different conditions sometimes prevail over different parts of these paths. In such cases the frequency will have to be lowered to satisfy conditions in the part of the path in which the critical frequency is lowest. Under such conditions, absorption is likely to occur in the part of the path with the higher critical frequencies. Normally the ionosphere does not vary greatly over a latitude range somewhat exceeding that of the United States.

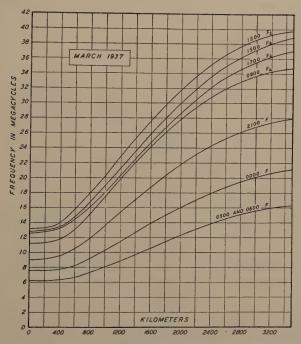


Fig. 12—Maximum usable frequencies, latitude of Washington, D. C., March, 1937.

The maximum possible distance for single-reflection transmission, corresponding to zero angle of departure, is about 2400 kilometers for E-layer transmission, and about 3500 to 4400 kilometers for F₂-layer transmission, depending on the virtual height of the layer. Practically, it is usually impossible to accomplish high-frequency transmission at this zero angle of departure over land because of absorption at the earth's surface. If a practical lower limit of $3\frac{1}{2}$ degrees is assumed for the angle of departure over land, the maximum distance for single-reflection transmission is about 1700 kilometers for E-layer transmission and 2800 to 3600 kilometers for F₂-layer transmission depending on the height. Single-reflection transmission may often be possible at

greater distances, while at the same time multireflection transmission over the same path may be more efficient.

IV. GRAPHS OF MAXIMUM USABLE FREQUENCIES

The maximum usable frequencies given in Figs. 1 to 14 represent monthly average conditions. The graphs for December represent winter conditions which hold for several months centered on the winter solstice, those for June represent summer conditions which hold for several months centered on the summer solstice, and those for March represent the spring and fall equinoctial conditions or transition con-

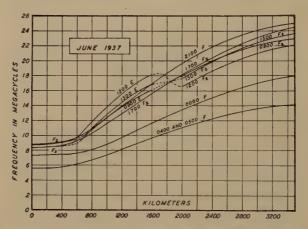


Fig. 13—Maximum usable frequencies, latitude of Washington, D. C., June, 1937.

ditions between winter and summer. The seasonal variations in characteristics of the ionosphere have been given in detail.

Day-to-day variations over short periods, such as a few weeks, are usually small (less than 10 per cent from the average) and have been given¹ in detail. Large variations, however, do occur during periods of ionosphere storms,⁴.⁵ which are usually associated with magnetic disturbances. At such times the virtual heights of the F₂ layer are increased and the critical frequencies decreased, resulting in lower maximum usable frequencies. The F₂ region becomes diffuse and unstable, resulting in high absorption and poor quality of transmission.

The graphs are given for single-reflection transmission at the latitude of Washington. Linear interpolations may be made between graphs

⁴ S. S. Kirby, T. R. Gilliland, E. B. Judson, and N. Smith, "The ionosphere, sunspots and magnetic storms," *Phys. Rev.*, vol. 48, p. 849; November 15, (1935).

⁵ S. S. Kirby, T. R. Gilliland, N. Smith, and S. E. Reymer, "The ionosphere, solar eclipse and magnetic storm," *Phys. Rev.*, vol. 50, p. 258; August 1, (1936).

for successive times. The conditions given in the curves are for local time at the geographical part of the ionosphere where the waves are reflected. For multireflection transmission the distance scale on the curves must be multiplied by the number of hops. For example, for two-hop transmission, 3000 kilometers on the graphs should be read as 6000 kilometers. For a path which does not involve latitudes differ-

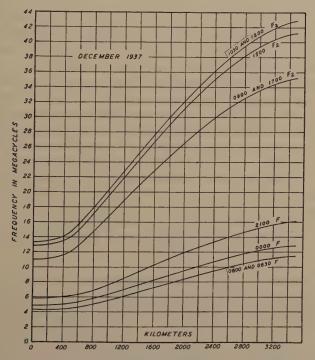


Fig. 14—Maximum usable frequencies, latitude of Washington, D. C., December, 1937.

ing widely from that of Washington, the maximum usable frequency is then the lowest one of the several corresponding to the local times at the several points of reflection along the path.

As an example of the use of the graphs, let us consider two-hop transmission from Washington to San Francisco, 4000 kilometers, at 0900, E.S.T., in December, 1937. The distance of each hop is 2000 kilometers, and the local times at the points of reflection are 0815 and 0645, at the eastern and western reflecting points, respectively. Interpolating between the graphs for successive times in Fig. 14 at the 2000-kilometer distance, we obtain the maximum usable frequency to be 21,000 kilocycles for the first or eastern hop, and 10,000 kilocycles

for the second or western hop. Thus 10,000 kilocycles is the highest frequency that can be used at this time for two-hop transmission over the given path.

The graphs for June show that both the E and F layers are effective. in determining the maximum usable frequencies. For example, at noon the F₂ layer will be effective out to a short distance, say 500 kilometers, because of its higher critical frequency; then the E layer will be effective out to 1600 kilometers because of its lower height and resulting larger angles of incidence. Beyond about 1600 kilometers the intensity of single-hop E-layer transmission decreases because of the necessarily small angles of departure at the transmitting station. The maximum usable frequency is then that which will be transmitted by two-hop E-, or single-hop F2-layer propagation. Either of these conditions gives as maximum usable frequency lower than does the single-hop E-layer transmission at about 1600 kilometers. Of the two conditions, the F₂: layer gives the higher maximum usable frequency for average conditions at Washington as indicated on the graphs. A similar decrease of maximum usable frequency occurs when the F-layer transmission must change from one-hop to two-hop propagation at about 3500 kilometers.

V. OTHER TRANSMISSION CHARACTERISTICS

1. Sporadic E

Reflections from the E region are often found at frequencies considerably in excess of the maximum daytime critical frequency for the E layer. These differ in character from the regular refraction phenomena of the normal E layer, and occur at all seasons, but are more prevalent during the summer months in Washington, sometimes persisting for several hours. Sporadic E reflections seem to be of the nature of reflections from a sharp boundary, as suggested by Kirby and Judson, at rather than of regular refraction from a region of intense ionization. Because of the sporadic nature of these reflections their effect has not been included in the graphs. Sporadic E, however, accounts for good transmission at higher frequencies than those indicated by the graphs for a small percentage of the time and at irregular intervals. Reports of long-distance transmission at frequencies as high as 56 megacycles, which occur mostly during the summer, appear to be by way of the sporadic E layer.

⁶ S. S. Kirby and E. B. Judson, "Recent studies of the ionosphere." Nat. Bur. Stand. Jour. Res., vol. 14, p. 469; April (1935); Proc. I.R.E., vol. 23, pp. 733-751; July, (1935).

2. Absorption

In contrast to the sharply defined maximum usable frequencies shown by the present graphs there are less well defined minimum usable frequencies for practical high-frequency sky-wave transmission for various distances. These lower-frequency limits are determined by absorption. The absorption varies in general with time of day, season, frequency, and length of path. Absorption occurs mainly in the lower ionosphere, that is, in the E layer and below. The absorption is greater during the summer day than during the winter day. It is greater for the lower high frequencies, that is, those frequencies which are close to or below the maximum usable frequency for the E layer for a given distance. It is usually greater during the day than during the night, especially for these lower high frequencies. A high-frequency wave at large angles of incidence corresponding to long distances is absorbed like a much lower-frequency wave at small angles of incidence, corresponding to short distances.

3. Scattered Reflections

Usually complex scattered reflections are observed at frequencies considerably higher than the F₂ critical frequencies.⁶ These reflections provide weak poor-quality transmissions much above the maximum usable frequencies for F-layer transmissions, over distances of several hundred kilometers.

VI. CONCLUSIONS

For practical applications the principal value of these graphs is to estimate transmission conditions in the future either diurnally, seasonally, or over longer periods. The ionosphere data,¹ from which these graphs were derived, indicate that the diurnal and seasonal characteristics are regular and may in general be predicted. In addition there is a large variation of ionization densities with the 11-year sunspot cycle as indicated by the increase of maximum usable frequencies from 1933 to 1937. At the end of the year 1937 the sunspot cycle was near a maximum and is expected to return to a minimum about 1944. In a rough way these graphs may be used for corresponding times on the descending part of the sunspot cycle. Current data are given by the National Bureau of Standards in its weekly ionosphere bulletins, broadcast from station WWV each Wednesday, and in reports on the "Characteristics of the Ionosphere at Washington, D.C.," published each month in the Proceedings.

USE OF FEEDBACK TO COMPENSATE FOR VACUUM-TUBE INPUT-CAPACITANCE VARIATIONS WITH **GRID BIAS***

By

R. L. FREEMAN

(Hazeltine Service Corporation, Bayside, Long Island, New York)

Summary—The input capacitance of amplifier tubes decreases with increasing grid bias, thereby detuning the input circuits of intermediate-frequency stages in a superheterodyne receiver. The phenomenon is caused by a shift in space charge and also by feedback through the grid-plate capacitance. The change in input capacitance is of the order of 1.5 micromicrofarads for each of the two causes. A simple form of feed-back coupling compensates for this change of capacitance. The grid-cathode capacitance and an unbypassed resistor in the cathode lead provide the feed-back coupling.

INTRODUCTION

HE input capacitance of vacuum tubes varies appreciably with changes in electrode potentials. The variation is caused by changes in space charge and in feedback through the grid-plate capacitance.

The capacitance variation causes a detuning of grid circuits in oscillators or in automatic-volume-controlled amplifiers. The detuning is serious for highly selective grid circuits, such as in the intermediatefrequency stages of sound receivers, or for grid circuits having very low capacitance, such as in the intermediate-frequency and radiofrequency stages of television receivers. This paper is confined to experimental and theoretical studies of intermediate-frequency amplifiers of sound receivers wherein compensation for input-capacitance variation with automatic-volume-control bias is obtained by the use of an unbypassed cathode resistor. However, the methods are applicable to radio-frequency and intermediate-frequency amplifiers in general.

Since completing this work, the writer has been informed by E. W. Herold of the latter's unpublished work and the publications of Strutt and van der Ziel. Both of those studies were primarly concerned with high-frequency amplifiers, and in them it was found that an unbypassed cathode resistor also reduces and stabilizes the input conductance component caused by transit-time effects at very high frequencies.

^{*} Decimal classification: R133×R142. Original manuscript received by the

York, N. Y., June 16, 1938. Presented before Thirteenth Annual Convention, New York, N. Y., June 16, 1938.

¹ M. J. O. Strutt and A. van der Ziel, "Simple circuit arrangement for improving the performance of high-frequency amplifier tubes in the short-wave band," Elek. Nach. Tech., vol. 13, pp. 260–268; August, (1936). See also British 18, 1415–1919. Patent No. 415,009.

MAGNITUDE OF DETUNING EFFECT

In Fig. 1 are shown resonance curves for a parallel-tuned circuit (f=450 kilocycles; C=100 micromicrofarads; Q=70) connected between the grid and the cathode of a 6K7 tube whose bias was varied. The plate and screen grid were bypassed to ground, and therefore

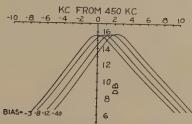


Fig. 1—Resonance curves for a tuned grid circuit showing the detuning caused by changes in space charge. Plate bypassed to ground.

the shift in resonance is attributable to the variation in the input capacitance caused by space charge. The maximum capacitance change is 1.1 micromicrofarads. An approximately linear relation was found between input capacitance and transconductance. Data obtained by more accurate methods by Jones² substantiate this.

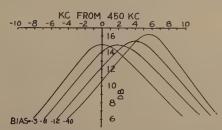


Fig. 2—Resonance curves showing the detuning caused by changes in space charge and in feedback through $C_{\it op}$. Intermediate-frequency transformer in plate circuit.

Fig. 2 shows resonance curves for the same tuned circuit when a typical double-tuned intermediate-frequency transformer, having 1.2-millihenry coils (C=100 micromicrofarads) and feeding a diode detector, was inserted in the plate circuit of the tube. The maximum shift in resonance is equivalent to a capacitance change of 2.6 micromicrofarads, of which 1.5 micromicrofarads is caused by feedback through C_{gp} in addition to the space-charge effect.

² T. Iorwerth Jones, "The dependence of the interelectrode capacitance of valves upon the operating conditions," *Jour. I.E.E.* (London), vol. 81, pp. 658-666; November, (1937). See Fig. 13.

The response curves for a complete intermediate-frequency stage from the modulator grid to the intermediate-frequency grid are shown in Fig. 3 for various values of bias applied to the intermediate-frequency tube. The transformers preceding and succeeding the intermediate-frequency tube (6K7) were of typical design having 1.2-millihenry coils (C=100 micromicrofarads). Inasmuch as higher-impedance transformers and tubes having greater values of C_{gp} are often used, the case is not an exaggerated one. When automatic-volume-control bias is used on such an amplifier, the detuning causes considerable asymmetry as well as loss in adjacent-channel selectance.

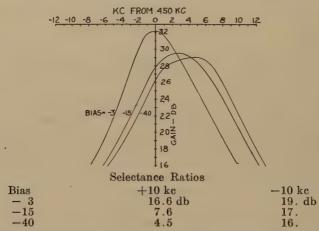


Fig. 3—Response of intermediate-frequency amplifier from modulator-tube grid to intermediate-frequency-tube grid for various bias voltages on the intermediate-frequency tube.

With very strong signals, the selectance ratio is relatively unimportant but audio-frequency distortion may then occur due to unequal transmission of the side bands and carrier. In the case of overcoupled transformers, an asymmetrical response curve results with large automatic-volume-control bias, and the over-all fidelity suffers.

COMPENSATING CIRCUIT

An unbypassed cathode resistor in combination with a capacitance from the grid to the cathode of a tube introduces feedback between the cathode and grid circuits which in effect reduces the capacitance between grid and ground with increasing transconductance, and thereby tends to compensate for the undesired variation.

In the basic circuit of Fig. 4, C_i represents the grid-to-cathode capacitance of the tube and any added capacitance that may be used

directly in shunt. It is sufficient to calculate the current I for an assumed voltage E to obtain the admittance Y_0 between grid and ground. For the present we shall center our attention upon the effective capacitance C_0 between grid and ground. To obtain a first approximation of C_0 we may neglect the effect of the plate-cathode capacitance C_{pk} and account for C_{qp} only partially by assuming the

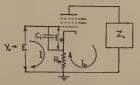


Fig. 4—Basic circuit for compensating for input-capacitance variation.

input capacitance caused by feedback through C_{gp} as being part of C_i . Hence we may write

$$E = I\left(\frac{1}{j\omega C_i} + R_k\right) - I_k R_k.$$

Since $I_k = -g_{1k}E_g = -g_{1k}I/j\omega C_i$, where g_{1k} is the transconductance between the grid and all other elements whose current flows through R_k , we obtain

$$Y_0 = rac{I}{E} = rac{j\omega C_i}{1 + g_{1k}R_k + j\omega C_i R_k} = G_0 + j\omega C_0.$$

Separating Y_0 into its real and imaginary components and neglecting $(\omega C_i R_k)^2$ in the denominator, we obtain for C_0

$$C_0 = \frac{C_i}{1 + g_{1k}R_k} {1}$$

The capacitance C_i is made up of three components: C_{gk} which is the value of capacitance between the grid and the cathode when the tube is biased to cutoff, C_s which is the increase in capacitance because of space charge, and C_f which is the input capacitance caused by feedback through C_{gp} . The latter two are experimentally proportional to g_{1k} .

Substituting $(C_{gk}+C_s+C_f)$ for C_i in (1) gives

$$C_0 = \frac{C_{gk} + C_s + C_f}{1 + g_{1k}R_k} {2}$$

 C_0 is made independent of g_{1k} and equal to C_{gk} when

$$R_k C_{gk} = \frac{(C_s + C_f)}{g_{1k}} \tag{3}$$

in which $(C_s + C_f)$ is evaluated at the corresponding value of g_{1k} .

Actually C_f is a function of frequency, as it is proportional to the real component of the plate-load impedance. The decrease in C_f away from resonance is unimportant as it only causes a very slight warping of the skirts of the grid-circuit resonance curve.

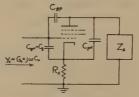


Fig. 5—Equivalent compensating circuit used for a detailed analysis.

The analysis leading to (1) neglects the plate-cathode capacitance C_{pk} and the effect of C_{gp} on the input conductance. A complete analysis based on the equivalent circuit of Fig. 5 shows that C_0 to a first approximation is fairly well expressed by (1). The conductance G_0 between grid and ground is expressed by the following approximate relation which neglects numerous terms which are negligible when C_{pk} is 2 micromicrofarads or less:

$$G_0 = \frac{\omega^2 (C_{gk} + C_s)^2 R_k}{(1 + g_{1k} R_k)^2} - \frac{g_{1k} X_L \omega C_{gp}}{1 + g_{1k} R_k} - \frac{g_{1k} R_k R_L \omega^2 C_{pk} (C_{gk} + C_s)}{(1 + g_{1k} R_k)^2} . \tag{4}$$

The quantity (R_L+jX_L) represents the plate-load impedance Z_L as measured between the plate and the cathode and which contains C_{pk} in parallel with (R_k+Z_2) , R_L is a maximum on or near resonance, while X_L is zero at resonance and reaches negative and positive maxima at frequencies respectively greater or less than resonance. The absolute maximum value of X_L is of the order of one half the maximum value of R_L .

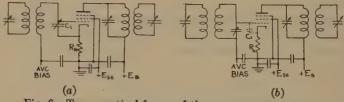


Fig. 6—Two practical forms of the compensating circuit.

Two practical forms of the compensating circuit are shown in Fig. 6. In Fig. 6(a) the trimmer condenser for the secondary of the intermediate-frequency transformer is returned to cathode, and thus C_{gk} is made large. This arrangement calls for a value of R_k of 10 to 20 ohms for a 6K7 tube. If self-bias is needed, a bypassed resistor may be added in series with R_k . In Fig. 6(b) the usual self-bias resistor is used

unbypassed for R_k , and the secondary trimmer and coil are returned to ground as is conventionally done. It is fortunate that the value of R_k which provides the exact compensation is very nearly the correct value for self-bias.

The second and third terms of the expression for the input conductance G_0 as expressed by (4) are negligible relative to the first for low values of transconductance. For a given pair of intermediate-frequency transformers both forms of compensating circuits require the same value for the product $R_k C_{gk}$. Since C_{gk} is much greater in the circuit of Fig. 6(a) the input conductance G_0 , as expressed by (4), will be much greater for low values of transconductance than for the circuit of Fig. 6(b) where it is, in fact, negligible.

When the transconductance is a maximum, G_0 may be made zero or even negative in the vicinity of resonance (which is the only region of importance) by returning the No. 2 grid and/or the No. 3 grid to the cathode. These optional connections change the value of C_{pk} and thus alter the magnitude of the negative third term of (4). The following ranges of values were obtained for C_{pk} for various samples of 6K7 metal tubes:

For the 6K7G glass tubes the minimum value of C_{pk} is 8 micromicrofarads for those makes in which the screen surrounding the plate is connected to the cathode. Because of this large capacitance, the 6K7G tube is unsatisfactory for use in the compensated circuit, the negative input conductance being too great. The older 6D6 glass tube does not have this outer screen, and therefore this tube should be satisfactory.

The presence of the unbypassed cathode resistor reduces the effective transconductance of the compensated tube by a factor of $1/(1+g_{1k}R_k)$. This factor is about 0.97 for the circuit of Fig. 6(a) but amounts to about 0.7 to 0.8 for the circuit of Fig. 6(b). The loss may be offset by making G_0 somewhat negative, thus raising the effective Q of the secondary tuned circuit.

The circuit used to obtain the curves of Fig. 2 was modified to conform with the arrangement of Fig. 6(a) by the addition of a 15-ohm

resistor for R_k , and the resonance curves of Fig. 7 were obtained. No attempt was made to minimize G_0 , as both the No. 2 and No. 3 grids of the 6K7 tube were returned to ground.

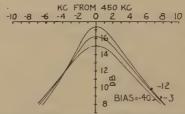


Fig. 7—Resonance curves for the same tuned circuit used for Fig. 2 except for the addition of a 15-ohm cathode resistor connected as in Fig. 6(a).

In Fig. 8 are shown grid-circuit resonance curves for the circuit of Fig. 9. The constants of this circuit differed from those of the one previously used in that the Q of the grid circuit coil was 115, and the

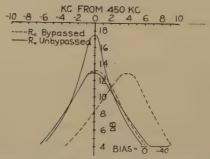


Fig. 8—Resonance curves for a tuned grid circuit with bypassed and unbypassed cathode bias resistor.

transformer in the plate circuit had a primary inductance of 0.6 millihenry (C=200 micromicrofarads). The exact circuit constants are given in Fig. 9. The resonance curves for the cathode resistor bypassed and

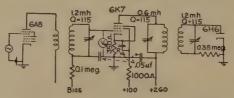


Fig. 9—Circuit used to obtain the curves of Fig. 8.

not bypassed are shown in Fig. 8. The greater resonant rise of voltage in the latter case was just sufficient to offset the degeneration in R_k so that the over-all gain from the 6A8 grid to the diode detector was the same for both cases.

BALANCED FEED-BACK AMPLIFIERS*

By

EDWARD L. GINZTON (University of California, Berkeley, California)

Summary—The following paper describes and analyzes mathematically a new principle in amplifier design. It is shown how noise, phase shift, frequency distortion, etc., can be largely eliminated by use of the "balanced feed-back principle" without reduction of the over-all gain of the amplifier.

By means of balanced feedback the linear range of an amplifier using ordinary receiving-type tubes has been extended from an upper limit of 600,000 up to 2,500,000 cycles. Theoretical and experimental results were found to check closely,

INTRODUCTION

N RECENT years an important change in amplifier design was introduced by the development of the stabilized feedback or the negative feed-back principle.1 This design represented a substantial improvement in the performance of amplifiers; by means of this principle the amplifier response can be made linear and practically independent of line-voltage fluctuations, and noise and distortions are reduced. These improvements are made possible by the return of a fraction of the output to the input to be used as a controlling voltage. Since the energy fed back is in opposite direction to the input signal, the feedback reduces the signal output, and in the cases where performance requirements are very strict, the reduction of the over-all gain may be so large that it becomes impractical to use such a system. It is desired, then, to develop an amplifier which will possess the advantages of the negative feedback without its disadvantage—the reduction of the over-all amplification. Such a design is possible by means of application of the balanced feed-back principle.

BALANCED FEED-BACK AMPLIFIERS

1. The Balanced Feed-Back Principle

The name balanced feedback is derived from the fact that two controlling voltages, balanced against each other, are used at the input of the amplifier for the purpose of controlling its performance. One of

¹ H. S. Black, "Stabilized feedback amplifiers," Bell Sys. Tech. Jour., vol. 13, p. 1; January, (1934).

^{*} Decimal classification: R132. Original manuscript received by the Institute, June 24, 1937; revised manuscript received by the Institute, September 20, 1937; revised manuscript received by the Institute, March 23, 1938. Presented before Pacific Coast Convention, Spokane, Washington, September 2, 1937.

these voltages is obtained from the output of the amplifier and is the conventional negative feedback which regulates the performance of the amplifier, and will be referred to in this paper as the negative feed-back voltage. The other feed-back voltage, in opposite polarity to the negative feedback, and hence called the positive feed-back voltage, is obtained in such a way that it is always proportional to the input signal and is independent of frequency over the range of frequencies it is desired to amplify linearly. The positive and negative feedbacks are made normally equal to each other, and therefore, when the performance of the amplifier is normal, they cancel each other and, hence, have no effect on the output. But if the output should for some reason differ

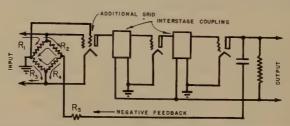


Fig. 1—An idealized balanced feed-back amplifier. An additional grid is used for the control of the performance.

from proportionality to the input, the feed-back circuits introduce voltages at the input of the amplifier in such a direction and of such a magnitude that the output is changed back to normal. In its operation, the balanced feed-back principle is very similar to the ordinary mechanical governor in steam engines, turbines, etc., in that the governor is acting only when the output differs from the desired normal operating conditions.

The principle of the balanced feedback is shown in the idealized circuit, Fig. 1. Except for the two feed-back circuits and additional control grid in the first tube, the amplifier is of the conventional design using whatever interstage coupling may be required for the specific application. Across the resistances of the bridge in the input circuit there are two voltages, one being the signal input, and the other being a fraction of the output returned to the input. Across the resistance R_3 the two voltages are in the opposite direction, as shown by current arrows in Fig. 1, and therefore, by proper choice of various resistances, the voltage across resistance R_3 can be made zero. As is seen in Fig. 1, resistance R_3 is connected between the cathode and the additional grid of the first tube; if the output is proportional to the input, the voltage across R_3 is zero and the additional grid has no effect on the

plate current of the first tube. But if the output deviates from proportionality, a voltage will be introduced across R_3 and the additional grid in the first tube will be at some potential with respect to its cathode. The output of the amplifier will be changed, accordingly, so as to tend to eliminate the original deviation and restore perfect proportionality.

2. Polarity of the Two Feedbacks

The feed-back voltages impressed at the input of the amplifier must satisfy two conditions. (1) The two feedbacks must be in opposite polarity. Since each tube and each interstage transformer, if they are used, produce a reversal of polarity, care must be taken to have the negative feed-back voltage actually of opposite polarity to the positive feedback. (2) The additional grid must be in proper polarity with respect to the control grid of the first tube. For example, if the amplifier output should become too low for some reason, then the voltage produced on the additional grid must be of the same polarity as the voltage of the control grid to increase the output to the proper level. Accordingly, the negative feedback must be of opposite polarity to the input and the positive feedback must be of the same polarity as the input.

These two conditions are obviously essential. Actual polarity of the two feedbacks depends upon the specific circuit used in the main amplifier and upon the manner in which the positive feedback is obtained.

3. Performance Equations for Balanced Feedback

The effect of balanced feedback as compared to the negative feedback is shown by the following derivation:

Let,

e =signal input voltage

A = amplification constant for the entire amplifier

Ae =signal output voltage without feedback

n = noise output voltage without feedback

d(E) =distortion output voltage without feedback

 β = propagation constant of the negative feed-back circuit, i.e., the fraction of the output returned to the input

 α = propagation constant of the positive feed-back circuit, i.e., a number times the input voltage. May be greater or less than 1.

E =signal output voltage with balanced feedback

N = noise output voltage with balanced feedback

D =distortion output voltage with balanced feedback.

The above notation is perfectly general and applies to an amplifier which utilizes both positive and negative feedback. All of the above symbols, except α and β , are obvious and have the usual meanings for all amplifiers. Definition of factor α , however, depends upon the particular amplifier and the way that positive feedback is introduced into the input. In general, α is defined by the ratio of the voltage developed by the positive feed-back circuit to the signal input voltage at the input grid of the first stage. Definition of β is perfectly general as defined above. For instance, in Fig. 1, assuming that R_1 , R_2 , R_3 , and R_4 are equal, $\beta = R_3/[2(R_3 + R_5)]$. Assuming that the two control grids in Fig. 1 have the same amplification factor, then $\alpha = R_3/[R_3 + R_1] = \frac{1}{2}$.

The output voltage without feedback is Ae+n+d(E). The balanced feed-back voltage is composed of positive feed-back voltage αe , and negative feed-back voltage $\beta(E+N+D)$. The total feed-back voltage is (at the input of the amplifier)

$$\alpha e + \beta (E + N + D). \tag{1}$$

The output voltage with balanced feedback is composed of the output voltage without balanced feedback plus the output due to the balanced feedback at the input. Hence, the output with the balanced feedback is

$$E + N + D = Ae + n + d(E) + \alpha Ae + \beta A(E + N + D) \quad (2)$$

$$(E + N + D)(1 - \beta A) = Ae(1 + \alpha) + n + d(E)$$
 (3)

$$E + N + D = \frac{Ae(1+\alpha)}{1-\beta A} + \frac{n}{1-\beta A} + \frac{d(E)}{1-\beta A} \tag{4}$$

If the amplifier is so adjusted that the positive feedback is equal and opposite to the negative feedback, i.e., $\alpha = -\beta A$, the performance equation for the balanced feed-back amplifier becomes

$$E + N + D = Ae + \frac{n}{1 - \beta A} + \frac{d(E)}{1 - \beta A}$$
 (5)

Equation (5) shows that distortion, noise, etc., are reduced by the factor $1/(1-\beta A)$, while the over-all gain of the amplifier is not affected by the introduction of the balanced feedback. Note that if the positive feedback is zero, $\alpha = 0$, and (4) becomes the performance equation for the conventional negative feed-back amplifiers.

$$E + N + D = \frac{Ae}{1 - \beta A} + \frac{n}{1 - \beta A} + \frac{d(E)}{1 - \beta A}.$$
 (6)

Equations (5) and (6) show the main difference between the balanced feedback and the negative feedback. In the negative feedback,

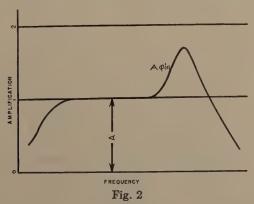
equation (6), the gain, noise, and distortion are all reduced by the same factor. In the balanced feed-back amplifiers there is a definite discrimination against noise and distortion, while the gain is not affected.

It must be remembered, however, that the only essential difference between (5) and (6) is the reduction of gain in the latter case. The same result can be achieved by means of negative feedback alone, if the amplifier possesses excessive gain, or if it is possible to increase sufficiently the mutual conductance of one or more tubes, or by addition of another stage. If this is possible the results obtained by means of negative feedback alone will be identical with those obtained by means of balanced feedback, with the additional advantage that the negative feed-back circuits tend to be simpler than those employed with balanced feedback. But as mentioned in the Introduction, there are cases where the particular requirements do not allow the reduction of the over-all gain of the amplifier.

From these equations the performance of the amplifier with balanced feedback may be determined. Frequency response, stability, delay distortion, critical points, etc., may be determined by proper manipulation.

4. Frequency Response of Balanced Feed-Back Amplifiers

The amplification constant of an amplifier is a complex quantity and is a function of frequency. The problem of calculation of the response of the amplifier can be subdivided for convenience into the calculation of the absolute value and the calculation of the corresponding



phase-shift angle. The two problems are very similar and, as can be shown mathematically, the phase shift and amplitude distortion take place simultaneously. In the following analysis, only the absolute value of the response will be considered, with the understanding that the

balanced feedback affects the phase shift in a similar manner and can be easily determined.

Referring to Fig. 2, let $\phi(f)$ be the factor by which the amplification of the balanced feed-back amplifier changes with frequency, $\phi'(f)$ be the factor by which the amplification constant A changes with frequency, and $\phi''(f)$ be the factor by which the positive feedback α changes with frequency. Assuming zero noise or other distortion, (4) becomes

$$E = A \phi'(f) e^{\frac{1 + \alpha \phi''(f)}{1 - \beta A \phi'(f)}}$$

$$= A e \phi(f)$$
(8)

$$= Ae\phi(f) \tag{8}$$

where,

$$\phi(f) = \phi'(f) \frac{1 + \alpha \phi''(f)}{1 - \beta A \phi'(f)}$$
 (9)

If
$$\alpha \phi''(f) \gg 1$$
 and $\beta A \phi'(f) \gg 1$, and $\alpha = -\beta A$. Then,

$$\phi(f) = \phi''(f). \tag{10}$$

This means that if the positive and the negative feed-back voltages are made normally equal and very much larger than one, the output will

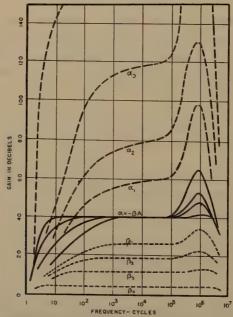


Fig. 3-Theoretical curves showing the effect of balanced feedback.

be a function of the performance of the positive feed-back voltage which can be made independent of frequency over the range of frequencies that it is desired to amplify without distortion, as is discussed later. If this condition is fulfilled, the performance of the entire amplifier will be independent of frequency over approximately the same range as the distortionless range of the positive feed-back circuit.

Fig. 3 shows the theoretical frequency-response curves obtained from (7) and a known typical curve for a two-stage resistance-capacitance-coupled amplifier. The dotted curves show the effect of the negative feedback alone; these show that the negative feedback makes the response of the amplifier linear over a large range of frequencies, but at the same time decreases the gain. The dashed curves show the effect of positive feedback alone. These are essentially the same as the response without any feedback, the difference being in the larger gain and the sharper cutoff outside the desired range. Balanced negative and positive voltages applied at the same time give a family of curves which approaches a straight line more and more as the ratio of the feedback to the input is made larger.

PRACTICAL BALANCED FEED-BACK AMPLIFIERS

1. Methods of Obtaining Positive Feedback

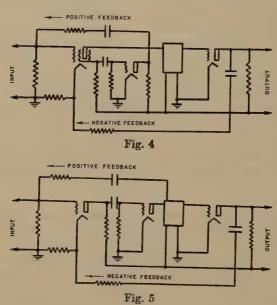
The above discussion and Fig. 3 show that it is desirable to have as high a value of balanced feedback as possible. There is no difficulty in obtaining a high value of negative feedback; the positive feedback is limited, however, by the type of circuit employed and tubes used. If a circuit similar to the one shown in Fig. 1 is used, and it is assumed that the additional grid has the same degree of control as the control grid, then the maximum value of positive feedback is one. This value is ordinarily too low, and either a more sensitive additional grid must be used or the positive feedback must be obtained in another way. At the present time there are no tubes on the market with two extremely sensitive grids, and therefore other methods must be devised to obtain the sufficiently high value of positive feedback.

Figs. 4 and 5 show possible methods of obtaining a high value of positive feedback. In Fig. 4, the second tube amplifies the output from the screen-grid circuit of the first stage. The output of the second tube is returned to the input of the amplifier as positive feedback, since the output of the second tube is in phase with the input of the amplifier. By proper adjustment of the circuit, the effective positive feedback can be made very large, as can be seen from (6) and from

$$E' = \frac{A'e'}{1 - \gamma A'} \tag{11}$$

where A' is now the amplification of the positive feed-back circuit and

 γ is its feed-back ratio. If $\gamma A'$ is made almost equal to 1, then the output becomes very large, thus making the effective input to the amplifier much higher. Fig. 5 shows a similar circuit, except that the positive feed-back circuit is also a part of the main amplifier.



In either of these two cases, the part of the amplifier that is involved in the positive feed-back circuit must be designed to give a linear response over the entire range of frequencies it is desired to amplify linearly. Since the gain of this part of the circuit does not have to be large, this requirement is usually not very difficult to fulfill in practice.

In all of the circuits considered, i.e., Figs. 1, 4, and 5, there are three voltages present at the input of the amplifier. One of these is the signal input voltage, the second is the conventional negative feed-back voltage obtained from the output of the amplifier, and the third is a voltage proportional to the signal input voltage and is used to balance out the negative feed-back voltage when the operation of the amplifier is perfect. This last voltage, termed above as the positive feed-back voltage, may be obtained in several ways mentioned above and shown in Figs. 1, 4, and 5. Its magnitude is always defined by its propagation constant α

$$\alpha = \frac{\text{positive feed-back voltage at input}}{\text{signal input voltage}}.$$
 (12)

The propagation constant of the negative feed-back circuit is defined as β and is always equal to

$$\beta = \frac{\text{negative feed-back voltage at input}}{\text{signal output voltage}}$$
 (13)

The positive feed-back voltage is determined by the type of the circuit used. For instance, in Fig. 1, the positive feed-back voltage is simply a fraction of the input voltage, depending upon the setting of the bridge. In circuits such as are shown in Figs. 4 and 5, the positive feedback is obtained by means of a regenerative network, the exact value of positive feed-back voltage being determined by means of (11).

If in the future a tube with a very sensitive additional grid, not necessarily linear, is developed, the circuit shown in Fig. 1 will become practical and ideal, since no additional parts are required to produce positive feedback.

2. Construction of Balanced Feed-Back Amplifiers

In actual construction of the amplifier used care was taken to minimize all stray capacitances between the feed-back leads and ground, as well as the stray capacitances between the various stages. By arranging the tubes horizontally in series, with shields in between, it was possible to make the stray capacitance of each stage small. By placing the tubes in a zigzag arrangement between two shields so that the input and the output of the amplifier are very close together, the feed-back leads are made very short and the feed-back circuits become practically independent of frequency.

DESIGN OF A BALANCED FEED-BACK AMPLIFIER

The general scheme of applying the balanced feed-back principle

to an amplifier of any type is as follows:

(1) Design of the high-gain amplifier so that the amplitude of the output is sufficiently high over the entire range of frequencies it is desired to amplify linearly. It is necessary to design it so that the amplification does not drop very far below the desired normal; it is unimportant how high it may become over certain ranges of frequencies. The design of the amplifier is conventional, the type of interstage coupling depending upon the specific requirements. In resistance-capacitance-coupled amplifiers the response curve, however, should be kept from dropping too low by means of small inductances placed in the plate leads of each tube, as suggested by Kell.² The value of the inductance

² R. D. Kell, "Description of experimental television transmitting apparatus," Proc. I.R.E., vol. 21, pp. 1674-1691; December, (1933).

depends upon the stray capacitance of each stage and the highest frequency desired to amplify.

- (2) The two feed-back circuits must be designed so that they are absolutely linear over the entire working range of frequencies. If both of the feed-back circuits are properly designed, the negative feedback will reduce the distortion and make the amplifier linear; the positive feedback will restore the gain to the former level, thus giving a high over-all gain without appreciable distortion of any kind.
- (3) The stability of the amplifier depends upon the performance of the positive and negative feed-back circuits. Providing the positive feed-back circuit is properly designed, that is, the regeneration is not allowed to become too great, it cannot cause any instability. If the main amplifier uses negative feedback across more than two stages, or in cases where transformer coupling is used, phase reversals may produce oscillations or "singing." However, H. S. Black1 has shown that negative feedback can be applied over as many as five stages without producing instability. Since the design of the negative feed-back circuits in a balanced feed-back amplifier in no way differs from the conventional negative feed-back amplifiers, the practical and theoretical limitations of balanced feed-back amplifiers are essentially the same as those of negative feed-back amplifiers. A simplified discussion of the question of stability, its dependence upon the phase shift of the ampli-

The main factors to consider in design of the feed-back circuits are:

fier, and the effect of feedback on phase distortion, will be found very well discussed in Black's article. A rigorous treatment of the question of stability is given by H. Nyquist,3 and E. Peterson, J. G. Kreer, and

- (a) The type of stabilization required, either current or voltage. If the negative feedback is proportional to the current output, the balanced feedback will tend to keep the current output constant, whereas, if the negative feedback is proportional to the output voltage, the latter will be stabilized.
- (b) Polarity. As previously stated, the negative feedback must be of opposite polarity to the input and the positive feedback must be of the same polarity as the input.

EXPERIMENTAL RESULTS

1. Restoration of Wave Shape

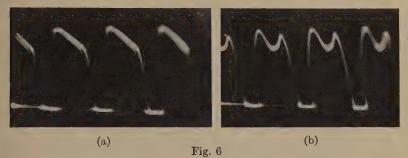
L. A. Ware.4

Figs. 6 and 7 show the effect of balanced feedback on the performance of an amplifier, with the tubes very badly overloaded, thus pro-

3 H. Nyquist, "Regeneration theory," Bell Sys. Tech. Jour., vol. 11, p. 126; July, (1932).

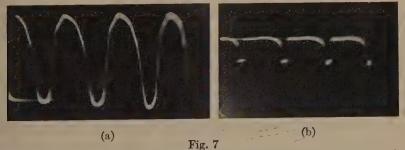
4 E. Peterson, J. G. Kreer, L. A. Ware, "Regeneration theory and experiment," Proc. I.R.E., vol. 22, pp. 1191-1210; October, (1934).

ducing distortion. Fig. 6(a) is an oscillogram showing the output without any feedback. Fig. 6(b) shows the effect of a small value of balanced



(a) An oscillogram showing the output of an amplifier with sine-wave input. Direct-current bias = -3 volts. Signal voltage is 10 volts.
 (b) An oscillogram showing the output of an amplifier under conditions of Fig. 6(a), but with a small value of balanced feedback.

feedback, and Fig. 7(a) shows the effect of a fairly large value of feedback. Although the wave shape is not completely restored to that of the sine-wave input, it is much improved. Figs. 6(a), 6(b), and 7(a) were taken under the same operating conditions, and show that the elimination of distortion took place without any reduction of the over-all gain. Fig. 7(b) is an oscillogram showing the feed-back voltage corresponding



(a) An oscillogram showing the output under the same conditions as Fig. 6, but with a large value of balanced feedback. While not a sine-wave, it resembles the sine-wave input much more than Fig. 6(a).

(b) An oscillogram showing the voltage across the feed-back bridge with the output as shown in Fig. 7(a).

to the feed-back conditions of Fig. 7(a); this is the voltage that must be amplified, subjected to distortion, and combined with the distorted output of Fig. 6(a) to produce the improved wave shape shown in Fig. 7(a). (The lopsidedness of the oscillograms, giving three values of voltage for one time in several places, is not due to the distortion in amplifier, but due to a faulty sweep of the cathode-ray oscillograph used to obtain these oscillograms.)

2. Frequency Characteristics of a Balanced Feed-Back Amplifier

Application of the balanced feed-back principle to a resistance-capacitance-coupled amplifier (two stages (Fig. 4)) improved its frequency-response characteristic as shown in Fig. 8. Curve A in Fig. 8 shows the response of the amplifier without any feedback and without

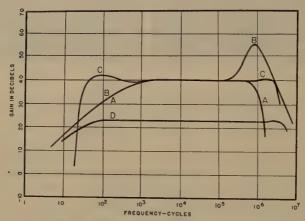


Fig. 8—Frequency response of an amplifier under various conditions. Curve C represents the response with balanced feedback, curve D with negative feedback

inductances in the plate leads. Curve B shows the response without feedback, but with inductances in the plate leads, which cause the large hump at the higher frequencies. The inductances are used, as described above, to prevent a too low cutoff frequency. Curve C shows the response with balanced feedback; the hump is practically entirely removed, and the regions of amplification that are not too low are restored to normal amplification. The same linear range can be obtained by means of negative feedback alone as shown by curve D, but only at reduced gain.

3. Degree of Predictability from Theoretical Considerations

Using the value of stray capacitance per stage, which depends upon the quality of construction, and the desired cutoff frequency, it is possible to calculate the performance of the amplifier. The response curve for the amplifier without feedback (this depends upon the type of coupling between stages, tubes used, stray capacitance, etc.) may be either calculated or obtained experimentally. The response curve for the amplifier with the balanced feedback may then be calculated from the theoretical considerations developed above, and by making use of (7). The accuracy with which these curves may be predetermined can be seen in Fig. 9. The theoretical curve, within the experimental errors in measurements, is the same as measured in the laboratory.

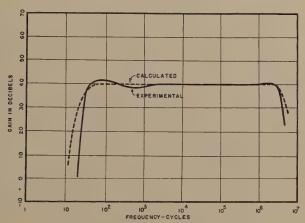


Fig. 9-Degree of predictability.

Conclusions

The experimental results verify the theoretical expectations and show that balanced feedback can be applied to various amplifiers, for various purposes, without undue elaborations. The theory agrees very closely with actual results, and therefore the design procedure is simple.

Balanced feed-back principle can readily be applied to wide-range television amplifiers. An amplifier using ordinary receiving-type tubes, i.e., RCA-57, has been built and its linear output was extended from a range of 1000 to 600,000 cycles to a range of essentially linear response from 30 to 2,500,000 cycles. By means of smaller tubes, such as RCA-954, it is to be expected that linear response up to 5,000,000 cycles can be obtained at a high gain.

Application of balanced feedback to audio-frequency amplifiers in radio sets may be of value. It would tend to eliminate noises that are developed in the last stage due to aging of the tube, gases, etc., and thus prolong the usable life of the tube.

....

OCCURRENCE OF "TWEEKS" ON A TELEPHONE LINE*

By

RUSH F. CHASE

(Madison Park Electric Company, Seattle, Washington)

N connection with some outdoor sound recording, the author had occasion to construct a telephone line, about 1500 feet in length, using a ground return. One end of the wire was temporarily earthed, while telephone receivers were connected to the opposite end of the line and the earth. In addition to the usual alternating-current hum, very faint "chirps" were heard at frequent intervals. Since these sounds were similar to those described briefly by Barkhausen,1 a further investigation was conducted.

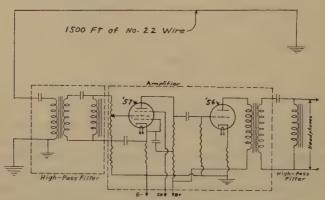


Fig. 1—Diagram of apparatus for observing "tweeks."

The diagram of the apparatus finally used to detect these sounds (commonly called "tweeks") is shown in Fig. 1. High-pass filters were used to good advantage in attenuating the interfering power-line noises. The over-all gain of the amplifier was approximately 65 decibels at 1550 cycles. Using the equipment shown, data were obtained involving the average number of tweeks occurring per minute as a function of the time at which observations were made. The results of one set of data are plotted as a curve in Fig. 4; in this case, observations were taken every hour, the number of tweeks being averaged over a five-minute period at each observation. These tweeks were apparently damped oscillations having a constant frequency in the neighborhood of 1550

tute, October 1, 1937.

1 Heinrich Barkhausen, "Whistling tones from the earth," Proc. I.R.E., vol. 18, pp. 1155-1159; July, (1930).

^{*} Decimal classification: R114. Original manuscript received by the Insti-

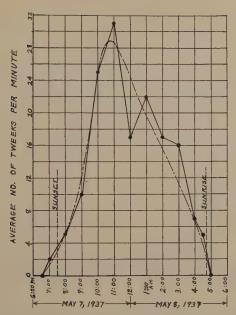


Fig. 2—Variation in occurrence of tweeks on a telephone line with ground return.

Length of line = 1200 feet open wire.

Dashed curve = average of plotted points.

Solid curve = actual graph of points.

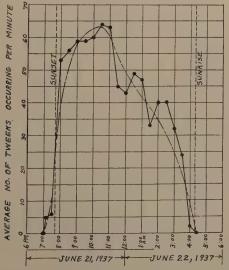


Fig. 3—Variation in occurrence of tweeks on a telephone line with ground return.

Length of line = 900 feet open wire.

Dashed curve = average of plotted points.

Solid curve = actual graph of points.

cycles, a value which was found to be independent of both line length and filter configuration. A few of these sounds were heard which had a frequency of 1800 cycles.

Similar observations were made next at an isolated location about thirty miles from Seattle, and a curve was plotted from data taken every half hour. This curve is given in Fig. 3, and is seen to be very similar to the one in Fig. 2.

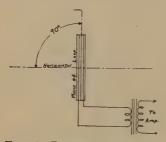


Fig. 4—Position of loop for maximum reception of tweeks.



Fig. 5—Position of loop for zero reception of tweeks.

Note that no tweeks were heard until shortly before sunset, and that they disappeared just as shortly after sunrise. Note also that a maximum was reached during the dark hours. Also, just before sunset and immediately after sunrise, the sounds were very highly damped oscillations of a frequency several hundred cycles above the normal value for dark hours. No tweeks were heard unless the line was grounded at the far end; when the ground return was thus used, the system was equivalent to a huge loop antenna with a vertical plane. To ascertain definitely the origin of the tweeks, a loop antenna was set up as illustrated in Figs. 4 and 5. When signals picked up on this loop were highly amplified, tweeks were plainly heard with the plane of the loop vertical, as in Fig. 4. However, when the plane of the loop was horizontal, Fig. 5, no tweeks were evident, indicating that the signals contained no horizontal electric-field component and the source of the sounds might be directly above or below the loop.

Barkhausen observed similar effects while experimenting with a submarine cable, and these are described by him in the article previously referred to. He suggests that the cause of the phenomenon is multiple reflection of a single static discharge, between the earth and some region of the ionosphere. On this basis the author of this paper makes the following mathematical analysis:

Referring to Fig. 6, assume that a static discharge occurs someplace in the atmosphere, say at S. The impulse travels outwards in all directions, and that portion of the wave front moving perpendicular to the earth's surface will strike the latter and be reflected vertically to the ionosphere. Reflection will take place at this point also, and the wave will again travel earthward to repeat the action. These reflections will continue until the energy of the static impulse is dissipated. Now, during this multiple-reflection process, the wave front must necessarily cut the wire on which observations are being made. In traveling the distance H_1 downward, shown in Fig. 6, assume a positive half cycle to be induced in the wire. When the reflected impulse returns over

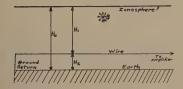


Fig. 6—Diagram showing path of static impulse.

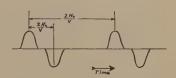


Fig. 7—Diagram illustrating induced voltages in line.

the distance H_2 , a half cycle of opposite or negative polarity will be induced. However, another half cycle of positive polarity will not occur again until the impulse has traveled a distance $2H_1$, or $2H_0$ (H_2 is negligibly small compared to H_1). Each successive peak will, of course, be lower than the preceding one. Now, let the velocity of the static impulse be V and designate the observed frequency of a single tweek² by f. Also, let f be the time of one cycle of frequency f. Then,

$$t = \frac{1}{f} \cdot \tag{1}$$

Also,

$$t = \frac{2H_1}{V} \cdot$$

Since $H_1 = H_0$ very nearly,

$$t = \frac{2H_0}{V} {2}$$

Equating (1) and (2), and solving for H_0 ,

$$H_0 = \frac{V}{2f} \,. \tag{3}$$

² The fundamental frequency component in Fig. 7 would be $V/2H_0$, as observed in the telephone receivers.

These values are illustrated in Fig. 7. Since V is 300,000 kilometers per second and f was found to be 1550 cycles by actual measurement, the computed height of the reflecting layer of the ionosphere is, by (3),

$$H_0 = \frac{300,000}{3100} = 99.4$$
 kilometers.

This value corresponds roughly with the actual height of the E layer, which is normally 110 kilometers at Washington, D.C. Bureau of Standards measurements regarding the height of the E layer in the vicinity of Seattle were not known to the writer.

These tweeks were observed on a cathode-ray oscilloscope, and a pictorial representation of the image seen on the screen is shown in Fig. 8. Owing to the short duration of the transients, satisfactory

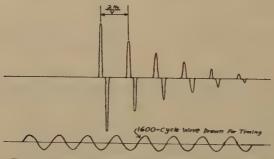


Fig. 8—Copy of cathode-ray picture of a tweek.

photographs could not be made. The 1600-cycle timing wave shown was, of course, sketched in afterward.

Conclusions

If these sounds are caused by multiple reflection from the ionosphere, some facts concerning the nature of the latter can be verified. First, since the frequency of the tweeks is always the same, except at twilight and at dawn, the height of the reflecting layer must be essentially constant for the location in which tests were made. Second, because the tweeks were heard only at night, and were most numerous near midnight, the reflecting properties of the layer must be best during the dark hours. The change in frequency and final disappearance of the tweeks at sunrise further verifies the effect of sunlight on the ionosphere.

Other related phenomena were noticed while obtaining data for this paper, and more information concerning this will be obtained at a later date.

THE TEMPERATURE COEFFICIENT OF INDUCTANCES FOR USE IN A VALVE GENERATOR*

By

E. B. MOULLIN

(Engineering Laboratory, University of Oxford, Oxford, England)

Summary.—This paper discusses the relative importance of skin effect on the temperature coefficient of inductance. It starts with a definition of inductance which includes explicitly the effect of magnetic field within the wire and analyzes inductance into an internal and external component. It is the internal inductance which has a temperature coefficient which can make the total coefficient greater than that of linear expansion. Its ability to affect the total coefficient obviously depends on the ratio of internal to total inductance. This ratio is assessed for various typical coils of round, solid wire, and found to be seldom less than about 5 per cent. It is also assessed for tubular and flat-strip conductors. The dependence of internal inductance on frequency is explored in general terms, and it is suggested that the product of highfrequency resistance and internal inductance tends to be constant for all conductors. An expression is derived for the ratio of the total temperature coefficient of a typical coil to that of linear expansion; it is shown that this ratio may well be four or five but that it need never exceed unity appreciably if the radius of the wire is chosen suitably with respect to the frequency. Some general rules are given for this choice.

Finally three methods of constructing coils on a ceramic former are compared in relation to the temperature coefficient.

I. Introduction

HE temperature coefficient of inductance is an important factor in radio communication, since it affects the frequency stability of valve generators. A thorough investigation has been made recently by H. A. Thomas¹ of the temperature coefficient of many coils of practical form.

In designing coils suitable for use in valve generators it is helpful to understand the various ways in which temperature operates to alter inductance. The writer² has already discussed several of these; J. Groszkowski³ also has pointed out other relevant factors quite recently. Since the whole problem is concerned with second-order differences, it is intricate in essence and demands very patient attention to some fundamental conceptions. It is hoped these notes may be helpful as a general guide in design of coils and for tracing the cause of large temperature coefficients. They contain nothing which has not been well known

^{*} Decimal classification:R145.3×R382. Original manuscript received by the

Institute, November 16, 1937.

¹ H. A. Thomas, Jour. I.E.E. (London), vol. 77, p. 702, (1935).

² E. B. Moullin, "The temperature coefficient of inductance, with special reference to the valve generator," Proc. I.R.E., vol. 23, pp. 65-84; January,

^{(1935).} ³ J. Groszkowski, "The temperature coefficient of inductance," Proc. I.R.E. vol. 25, pp. 448-464; April, (1937).

for about a century; but however well known in principle, the information does not seem to have been marshalled to bear on this particular problem.

II. INTERNAL AND EXTERNAL INDUCTANCE

The current in any coil produces a magnetic field which is a store of energy. If the stored energy of the whole magnetic field be W when the current in the coil is i, it is usual to write

$$W = \frac{1}{2}Li^2. \tag{1}$$

The factor L in this equation depends only on the geometry of the coil and is called its self-inductance. Equation (1) is the simplest definition of self-inductance which is rigorous. It is very common to define self-inductance in terms of flux linkage through the circuit; though this definition is correct, it requires a bulky parenthetic definition of the word circuit. For any wire has a finite cross section, and thus there are an infinite number of closed contours which lie everywhere inside the wire; which of these is to be regarded as the circuit, 4,5 since its boundary is not precise? If a fine-wire coil is wound on a highly permeable iron core it is pedantic to discuss the exact meaning of the circuit. But conditions are very different for the coils used in a valve generator and a vaguely stated definition by flux linkage is competent to lead to gross errors in estimating temperature coefficients.

Let us here define self-inductance in terms of energy and see how to calculate it. If at any given point of space there is a magnetic field H, then it can be shown that the energy of the whole field may be calculated as though it were distributed at the rate of $H^2/8\pi$ per unit volume. In symbols this becomes

$$W = \frac{1}{8\pi} \iiint (H_{1}^{2} + H_{2}^{2} + H_{3}^{2}) dx \cdot dy \cdot dz, \tag{2}$$

the integration being performed throughout all space. The words "as though" should be noted: there is no evidence to show where or how the energy is stored. Equation (2) is a mathematical tool which will yield the correct answer if no portion of space is omitted from the summation.

The magnetic field of a coil fills all space and exists inside the wire as well as outside it. Hence the flux inside the wire contributes to L and the contribution is always finite, however fine the wire. It should

Press, (1916).

⁴ The writer has discussed this in detail on p. 328 of his "Radio Frequency Measurements," second edition, Charles Griffin and Company, (1931), and has there laid down a valid interpretation of *circuit*.

⁵ See also F. B. Pidduck, "A Treatise on Electricity," p. 255, Cambridge

now be clear from (2) that the energy definition does not involve uncertainties in respect of the finite volume of the wire.

The magnetic field outside the wire is independent of the frequency of the current (at least at all points whose distance is a small fraction of a wavelength). On the other hand the magnetic field inside the wire depends very much on the frequency and in a manner depending on the specific resistance and diameter of the wire. It falls asymptotically to zero as the frequency tends to infinity and in general is negligible at a frequency well inside the radio spectrum. Thus to evaluate (2) we have to add the contributions from two regions: in one region the contribution depends on frequency and in the other it does not. The possible effect of frequency depends on the relative contribution from the two regions at zero frequency. The external contribution has no temperature coefficient save that due to the change in dimensions of the coil. The internal contribution has a temperature coefficient which is a function of the temperature coefficient of resistance of the wire. Since the temperature coefficient of resistance is about 4000 parts in 106 per degree centigrade and the coefficient of linear expansion is about 17 parts in 106, it is at once obvious that the temperature coefficient of inductance can be very much greater than that of linear expansion, if the internal contribution is appreciable. However, the internal contribution can always be made negligible by a suitable choice of wire diameter. It is important it should be negligible in coils whose temperature coefficient is to be as small as possible. There is seldom any practical difficulty in making it negligible. This was discussed previously,6 and the effect dismissed as unimportant. Since, however, there is some indication that coils are still made, with the hope of small temperature coefficient, in which the effect of the internal contribution appears to have been overlooked, the ensuing discussion may be of practical use.

III. THE INTERNAL INDUCTANCE IN CERTAIN SIMPLE CASES

Since the internal inductance may have a large temperature coefficient, it may be useful to assess its relative magnitude in certain simple cases.

(A). Internal Inductance of an Isolated Straight Wire

The magnetic field inside a round wire bearing a uniform current density increases uniformly as the distance from the center. Hence

$$W = \frac{1}{8\pi} \int_{0}^{R} 2\pi r H^{2} dr = \frac{I^{2}}{R^{4}} \int_{0}^{R} r^{3} dr = \frac{I^{2}}{4}$$
 (3)

⁶ See page 69 of footnote reference 2.

Hence, from (1), $L_i = \frac{1}{2}$ per unit length of wire. It is important to notice that this is independent of the radius of the wire and therefore does not vanish with it.

The external inductance of a long, isolated wire can have no meaning: it is essential to define where the return wire is. The field inside any element of wire must depend on the shape of the circuit: it can always be expressed as the sum of the field used in (3), together with a superposed field, which we may choose to attribute to the remainder of the circuit. Such a component will be termed the penetrating field and we must find the energy contributed by it.

(B). Internal Inductance Caused by a Uniform Penetrating Field

Suppose a round wire is placed perpendicular to a uniform magnetic field H_p . If there is no other field inside the wire, then H_p inside the wire will contribute to the summation an amount $W_p = R^2 H_p^2/8$. If H_p exists in addition to the internal field of the current in the wire, then brief consideration will show that the resultant field inside the wire makes a contribution to the total energy which is the sum of the contributions made by the two component fields separately in the absence of the other. Hence the internal inductance of any short length δl of wire in any circuit is

$$L_i = 2(\frac{1}{4} + R^2 H_p^2 / 8I^2) \delta l, \tag{4}$$

provided always that H_p is sensibly uniform over the cross section, It is not necessary to complicate this discussion by considering cases in which it is not uniform.

(C). Two Parallel Round Wires, Go and Return

Consider two long parallel wires, each of radius R, separated by a large distance D center to center. Each wire is situated in a sensibly uniform field $H_p = 2I/D$. The external inductance is well known to be

$$L_0 = 4 \log \frac{D-R}{R}$$
, per unit length.

Hence, by (3) and (4),

$$L_0 + L_i \equiv L_1 = 4\left(\log\frac{D - R}{R} + \frac{1}{4} + \frac{R^2}{2D^2}\right) \text{ per unit length}$$

$$\stackrel{.}{=} 4\left(\log\frac{D}{R} - \frac{R}{D} + \frac{1}{4} + \frac{3R^2}{2D^2}\right), \text{ if } D \gg R$$

$$\stackrel{.}{=} 4\left(\log\frac{D}{R} - \frac{R}{D} + \frac{1}{4}\right). \tag{5}$$

⁷ See, for example, National Bureau of Standards Bulletin No. 74, p. 245.

Equation (5) shows that the effect of the penetrating field is always a second-order quantity.

As the frequency tends to infinity, the field inside the wire tends to vanish: the exact limiting value of inductance is well known⁸ and is

$$L_{2} = 4 \log \frac{D + \sqrt{D^{2} - 4R^{2}}}{2R}$$

$$\stackrel{.}{=} 4 \left(\log \frac{D}{R} - \frac{2R^{2}}{D^{2}} \right) \text{ if } D \gg R$$

$$\stackrel{.}{=} 4 \log \frac{D}{R} \qquad (6)$$

Writing $L_1 - L_2 = \Delta L$, it follows from (5) and (6) that

$$\frac{\Delta L}{L_2} \doteq \frac{1}{L_2}$$
.

If D/R = 10, $L_2 = 9.2$ centimeter per centimeter,

if D/R = 100, $L_2 = 18.4$ centimeter per centimeter.

Hence in practice $\Delta L/L_2$ is likely to be of the order of, say, 7 per cent.

(D). Circle of One Turn

It can be shown⁹ that the magnetic field perpendicular to the plane of a circle of radius a and at a small distance c from the center of the wire of radius R is

$$H = 2I\left(\frac{1}{2a} \pm \frac{1}{c}\right). \tag{7}$$

Consideration of (7) will show that we may write

$$H_p = \frac{2I}{2a}$$

and hence that the contribution to the total energy from the field inside the wire may be written

$$W_1 + W_2 = \frac{1}{4} \left(1 + \frac{R^2}{2a^2} \right) I^2$$
 per unit length.

It is well known¹⁰ that the contribution from the field external to the wire is

10 See page 250 of footnote reference 7.

<sup>See page 31 of footnote reference 4.
See, for example, E. B. Moullen, Principles of Electromagnetism," p. 31.
Clarendon Press, (1932).</sup>

$$W_3 = \frac{1}{2} \times 4\pi a \left(\log \frac{8a}{R} - 2 \right) I^2.$$

Hence,

$$L_1 = 4\pi a \left(\log \frac{8a}{R} - 2 + \frac{1}{4} + \frac{R^2}{8a^2} \right).$$

Here again the effect of H_p is negligible if R/a is less than, say, 1/5. It follows that

$$\frac{\Delta L}{L_1} = \frac{\frac{1}{4} + R^2/8a^2}{(\log 8a/R - 2 + \frac{1}{4} + R^2/8a^2)}$$

$$= 9.4\%, \quad \text{if} \quad a/R = 10,$$

$$= 5\%, \quad \text{if} \quad a/R = 100.$$

or

Hence, once more, $\Delta L/L_1$ is of the order of, say, 7 per cent.

(E). Long Solenoidal Current Sheet

Let the solenoid have an inner radius R and a wall thickness t. If the current density in the wall is uniform, the field strength in the wall is proportional to the distance from the outer surface: in the limit of infinite frequency, this field is zero and the current is concentrated at the inner radius. Straightforward integration shows that

$$\frac{\Delta L}{L_2} = \frac{2}{3} \frac{t}{R} \left(1 + \frac{3}{2} \frac{t}{R} \right). \tag{9}$$

If t/R = 1/10, $\Delta L/L_2$ is once more about 7 per cent.

(F). Single-Layer Solenoid of Fine Wire

No attempt will be made to evaluate the contribution due to the penetrating field. Let the solenoid have radius a, length b, and N turns. Then if $b/a \le 1/4$, we may write¹¹

$$L_0 \doteqdot \frac{(2\pi aN)^2}{0.868a+b},$$

hence, from (3),

$$L_1 \doteq \frac{(2\pi aN)^2}{0.868a+b} + \frac{2\pi aN}{2},$$

¹¹ This follows because the factor K (see National Bureau of Standards Bulletin No. 74, pp. 252 and 253) can be represented very accurately by a hyperbolic equation, provided $b/a > \frac{1}{4}$.

therefore,

or

$$\frac{\Delta L}{L_1} = \frac{1}{\frac{4\pi aN}{0.868a + b} + 1}$$

$$= \frac{1}{4\pi N + 1}, \quad \text{if} \quad b/a = 0.23$$

$$= \frac{1}{N + 1} \quad \text{if} \quad b/a = 12.$$

This suggests that $\Delta L/L_1$ is of the order of 1 per cent, but (10) is necessarily an underestimate.

Examples (A) to (F) show that in cases of practical interest $\Delta L/L_1$ is of the order of, say, 5 per cent. Thus at any frequency a sufficient increase of temperature could cause some 5 per cent increase of inductance, even though physical expansion was prohibited. Since the internal inductance is necessarily an appreciable fraction of the total, it is competent to make the temperature coefficient much greater than that of linear expansion. Evidently it is wise to make the internal inductance negligible; if possible by a wise choice of wire diameter in relation to the working frequency and/or by making the internal inductance inherently small.

IV. Constructional Method of Making Internal Inductance Small

(A). Tubular Construction

If a solid round wire is replaced by a tube whose outer and inner radius are R_0 and R_i , respectively, the field within the metal for unit current is

$$H = \frac{2}{r} \; \frac{r^2 - R_i^2}{R_0^2 - R_i^2} \; \cdot \;$$

It follows after much reduction of straightforward integration that

$$L_{i} = \frac{1}{6} \frac{t}{R_{i}} \left(1 - \frac{5}{2} \frac{t}{R} + \frac{18}{5} \frac{t^{2}}{R^{2}} - \frac{139}{40} \frac{t^{3}}{R_{i}^{3}} + \cdots \right)$$
(11)

provided that $t/R_i \ll 1$, where $t = R_0 - R_i$.

(Note: The general expression shows that when $t/R_i = 1$, then $L_i = 0.322$ centimeter per unit length, vice 0.5 centimeter for a solid wire.)

Equation (11) shows that when the tube is thin, L_i becomes very small compared with its value for a solid wire: thus if $t/R_1=1/10$, L_i

is reduced in the ratio 1/30; by using tubular conductors the internal inductance can be reduced until the effect of the penetrating field becomes dominant. Thus for two parallel thin tubes, as shown in: Section III (B),

 $L_i = 2\left(\frac{1}{6} \frac{t}{R} + \frac{R^2}{D^2}\right)$ per unit length. (12)

The two terms in (12) are equal when $t/R = 6R^2/D^2$ and in practice this condition is scarcely likely to be met. It would seem, however, that the use of tubular conductors whould probably reduce $\Delta L/L$, from about 5 per cent to less than $\frac{1}{2}$ per cent.

(B). Flat-Strip Conductors

Consider a flat strip of thickness t and breadth b, great compared with t. The magnetic field inside the strip, due to the current in it, increases from zero at the mid-plane to $2\pi I/b$ at the surface. This supposes that $b/t \gg 1$ and that consequently end effects may be ignored. Accordingly simple integration gives

$$L_i = \frac{\pi}{3} \times \frac{t}{b} \text{ per unit length.}$$
 (13)

If the conductor had been composed of N circular wires of diameter t, placed side by side to make a band of width b=Nt, then it can be shown that

Equation (13a) is an underestimate since the penetrating field of neighboring wires is ignored.

(C). Single Turn of Flat Strip

The external inductance of a single-layer solenoid having N close turns in series is 12

$$L_0 = 2K \frac{a^2 N^2}{b},$$

where K is a parameter depending on b/a. If the N turns are in parallel, this becomes

$$L_0 = \frac{2Ka^2}{b}$$

= 40a, if $b/a = 1/5$.

¹² See page 252 of footnote reference 7.

Then we may write

$$L_1 = 40a + \frac{2\pi^2 a}{3} \frac{t}{b}, \text{ by (13)}$$
$$= 40a \left(1 + \frac{1}{6} \frac{t}{b}\right), \tag{14}$$

and then

$$\frac{\Delta L}{L_1} = 3\%$$
, if $t/b = 1/5$.

Compare this with a single turn of round wire, whose perimeter equals that of the strip. Then at high frequencies the two circles will have about the same resistance. This condition requires that $(12/5)b = 2\pi R$, hence $R = (6/25\pi)a$. It may then be calculated from (8) that $L_2 = 33.4a$ and $\Delta L/L_1 = 8.5$ per cent. The advantage of using flat strip is thus apparent: the advantage, however, is less than can readily be obtained by using tubes. Be it noted, however, that the flat strip must not be used on edge, for this will enhance $\Delta L/L_1$ since at high frequencies the current will concentrate at the inner radius.

Equation (9) may be used to form a rough estimate of this effect, and it follows that $\Delta L/L_1$ will be of the order of 2t/3a, and this might well be of the order of 1/4 without resorting to an exaggerated construction.

V. INTERNAL INDUCTANCE AS A FUNCTION OF FREQUENCY

(A). Solid Conductors

Having estimated the order of $\Delta L/L_1$ in various representative cases, we will now see how ΔL tends to zero with frequency, and consider first the isolated round wire.

The effective resistance R_n and internal inductance L_{in} of a long, round wire of radius R both depend on a parameter¹³ Z, such that

$$Z^2 = \frac{8\pi^2 nR^2}{\rho} \cdot$$

 R_n and L_{in} are expressible in terms of Bessel functions, which are more convenient to handle when expressed in polar form. They can also be expressed as power series in Z: it is always necessary to have two series,

¹³ An expression for R_n and L_{in} in polar Bessel form will be found on p. 140, equations (22) and (23), of "Bessel Functions for Engineers," by N. W. McLachlan, Oxford University Press, (1934), and appropriate tables on pages 182 and 183.

one for Z small and one for Z large. In the range of Z between 2.5 and 4, both series are troublesome, and it is necessary to use the Bessel tables. These series are

for Z small,

$$R_n/R_0 = \left(1 + \frac{Z^4}{192}\right),\tag{15}$$

and

$$L_{in} = \frac{1}{2} \left(1 - \frac{Z^4}{384} \right). \tag{16}$$

Equations (15) and (16) are correct to within 2 per cent up to Z=2.5; for Z larger than 2.5 the accuracy decreases rapidly.

For Z large

$$R_n/R_0 = \frac{Z}{2\sqrt{2}} \left(1 + \frac{1}{Z\sqrt{2}} + \frac{3}{8Z^2} + \frac{0}{Z^3} - \frac{1}{2Z^4} + \cdots \right)$$
 (17)

and

$$L_{in} = \frac{\sqrt{2}}{Z} \left(1 - \frac{3}{8Z^2} - \frac{3}{4\sqrt{2}Z^3} - \dots \right). \tag{18}$$

Equations (15) to (18) disclose the interesting fact that $R_n \times L_{in}$ is very nearly constant. Thus,

for Z small,

$$R_n \times L_{in} = \frac{1}{2} R_0 \left(1 + \frac{Z^4}{384} + \cdots \right)$$
 (19)

for Z large,

$$R_n \times L_{in} = \frac{1}{2} R_0 \left(1 + \frac{1}{Z\sqrt{2}} - \frac{9}{8\sqrt{2}Z^3} \cdots \right).$$
 (20)

Thus $R_n \times L_{in} = L_{i0} \times R_0 = 1/2R_0$ to an accuracy within 1 per cent if Z is less than 1.4 or greater than 70. Plotting from the tables shows that this product rises to a maximum when Z = 4 and then $R_n \times L_{in} = 1.1516$ in $R_0 L_{i0}$.

Some particular values are shown in Table I.

TABLE I

Z	0	1.5	2	2.5	3	4	5	6	10	20	40
$\frac{R_n L_{in}}{R_0 L_0}$	1	1.012	1.036	1.074	1.14	1.152	1.136	1.13	1.07	1.035	1.018
L_{in}	0.5	0.491	0.48	0.455	0.433	0.344	0.278	0.235	0.141	0.070	0.035

That $R_n \times L_{in}$ is sensibly constant is very suggestive and probably it is substantially true for solid conductors of any cross section, possibly inclusive of the effect of a penetrating field. It is precisely constant for a semi-infinite flat plate: for such it is well-known¹⁴ that $R_n = \rho m$ and $L_{in} = 2\pi/m$, both per unit width of plate, where m is a parameter analogous to Z, and thus $R_n \times L_{in} = 2\pi \rho$. This result may be applied safely¹⁵ to any solid cross section, provided the penetration depth 1/m is small compared with the radius of curvature. For copper at 15 degrees centigrade

$$\frac{1}{m} = \frac{6.55}{\sqrt{n}} \text{ centimeters} \tag{21}$$

Hence is $n=10^6$ cycles, 1/m=6.55/100 millimeters. Thus for frequencies not less than 1 megacycle, we may consider $R_n \times L_{in} = 2\pi \rho$ at all points of a surface whose radius of curvature is greater than, say, 1 millimeter: in particular for round wires not smaller than standard wire gage 18.

(B). Tubular Conductors

The exact expression for L_{in} of a tube can be stated in terms of Bessel functions, but is too cumbersome to present in any simple form. We can, however, infer its value as follows. For a solid wire, 15

$$L_{in} \doteq \frac{\sqrt{2}}{Z}$$

$$= \frac{1}{R} \sqrt{\frac{\rho}{8\pi^2 n}}$$

$$\equiv \frac{1}{R} \sqrt{\frac{\rho}{8\pi^2 n}}$$
(22)

where 1/m is the penetration depth described in (21). For a steady current in a tube,

$$L_{i0} = \frac{1}{6} \frac{t}{R} {.} {(11)}$$

Hence if 1/m is greater than, say, t/2, we may conclude that L_i has

¹⁴ See page 241 of footnote reference 4.

¹⁵ The comparison at the end of Section IV(C) may appear to be discordant with the relation $R_n \times L_{in} = \text{constant}$, for if both have the same R_n both should have the same L_{i0} , which is certainly not true. The discrepancy is due mainly to the sweeping assumption of equality of R_n by virtue of equal perimeters. The field penetrating the strip would enhance R_n and presumably by an amount such as to fulfill approximately the relation $R_n L_{in} = \text{constant}$.

changed little from its steady-current value and if $1/m \ll t/6$, we may calculate L_{in} from (22). Thus for copper tubes with walls 0.5 millimeter thick, (11) will be valid for frequencies up to about 40 kilocycles and (22) will be valid for frequencies higher than about 1 megacycle.

(C). Field Penetrating a Tube

If a tube is perpendicular to a uniform magnetic field which pulsates in simple harmonic fashion, it can be shown¹⁶ that the internal field remains uniform, but is reduced to a value H_p , where

$$\left| \frac{H_p'}{H_p} \right|^2 = \frac{1}{1 + x^2},\tag{23}$$

and

$$x = \frac{4\pi Rnt}{\rho} \cdot$$

If t=1/2 millimeter, R=5 millimeters, and n=5 kilocycles, then for copper x=3, and hence by (23), the square of the penetrating field is reduced to 10 per cent of the square of its external value. Hence, for such tubes the effect of penetrating fields may be ignored at radio frequencies.

VI. TEMPERATURE COEFFICIENT AS A FUNCTION OF FREQUENCY (A). Single-Turn Circle

We may write

$$L_n = 4\pi a(X + \frac{1}{2}L_{in}),$$

where $X = (\log 8a/R - 2)$ and L_{in} is given by (18).

Therefore,
$$\frac{1}{4\pi} \frac{dL_n}{dT} = \left(X + \frac{1}{2} L_{in}\right) \frac{da}{dT} + \frac{1}{2} a \frac{dL_{in}}{dZ} \cdot \frac{dZ}{dT}$$

Let α be the temperature coefficient of linear expansion and β that of resistance. Then,

$$\frac{da}{dT} = \alpha a$$
 and $\frac{dZ}{dT} = \left(\alpha - \frac{1}{2}\beta\right)Z$,

therefore,

$$\frac{1}{4\pi} \frac{dL_n}{dT} = a \left\{ \alpha \left(X + \frac{1}{2} L_{in} \right) + \frac{Z}{2} \left(\alpha - \frac{1}{2} \beta \right) \frac{dL_{in}}{dZ} \right\},\,$$

¹⁶ E. B. Moullin, "Principles of Electromagnetism," p. 94, Oxford University Press, (1932). However, equation (23) is valid only so long as $t \ll 1/m$: this restriction, however, does not invalidate the previous statement.

therefore,

$$\frac{dL_n}{dT} = \alpha L_n \left\{ 1 + \frac{Z}{2X} \left(1 - \beta/2\alpha \right) \frac{dL_{in}}{dZ} \right\}. \tag{24}$$

It follows from (19) that

$$Z \frac{dL_{in}}{dZ} = -\frac{\sqrt{2}}{Z} \left(1 - \frac{9}{8Z^2} - \frac{3}{\sqrt{2} Z^3} - \cdots \right)$$

$$\stackrel{.}{=} -L_{in} \left(1 - \frac{5}{8Z^2} \right), \text{ valid for } Z > 4, \text{ i.e., } L_{in} < 0.344,$$

$$\stackrel{.}{=} -L_{in}.$$

Hence the resultant temperature coefficient is

$$\alpha' = \alpha \left\{ 1 + \frac{L_{in}}{2X} \left(\frac{\beta}{2\alpha} - 1 \right) \right\}.$$

For copper, $\beta/\alpha = (4 \times 10^3/17)$, and hence $\beta/2\alpha = 118$,

therefore,
$$\alpha' = \alpha \left(1 + \frac{58L_{in}}{X} \right)$$
. (25)

It was estimated in Section II (D) that the greatest value of L_i/X was in practice likely to be about 7 per cent. Hence by (25), α'/α must be less than about 5. Equation (25) is, through (18), valid only for Z > 3, and indeed when Z is small α'/α tends to unity: closer examination shows it is a maximum near Z = 4, when $L_{in} = 0.344$. It follows from (18) that

$$L_{in} \doteq \frac{2\sqrt{2} L_{i0}}{Z},$$

and hence

$$\alpha'/\alpha = \left(1 + \frac{162}{Z} \frac{L_{i0}}{X}\right). \tag{26}$$

Hence if L_i/X is of the order of 7 per cent, then α'/α is less than 5/4 provided that Z is greater than 45. This requires that the wire shall be a standard wire gage 10 if the frequency is 1 megacycle and standard wire gage 16 if it is 10 megacycles. Although (26) was derived for a single-turn circle, it will be substantially correct for all coils having only a few well-spaced turns.

The maximum possible value of α'/α will be much reduced by using a tubular conductor, but the tube and solid wire will be indistinguishable for large values of Z.

We may perhaps summarize these findings as follows. For frequencies of the order of 10 megacycles, α'/α should not be more than about $1\cdot 15$ provided that the solid wire is a standard wire gage 10 and the coil diameter greater than 4 centimeters. For lower frequencies the wire must be larger than standard wire gage 10 and should preferably be tubular. (It is perhaps well to remind the reader that the specific resistance of commercial copper tubing is often twice that of electrolytic copper.)

VII. THE USE OF CERAMIC FORMERS

Nonexpanding ceramic formers are sometimes used in the attempt to produce coils having a small temperature coefficient. Sometimes the wire is wound in a screw thread cut in the former: then it is perhaps natural to use a wire not larger than, say, standard wire gage 20. For this wire Z is 3.8 at 150 kilocycles and 31 at 10 megacycles; at the first frequency α'/α will be near its maximum and at the second L_{in} is 0.05 (vice 0.5 for zero frequency). Estimating $\Delta L/L$ at 3 per cent (see (10)), it appears from (25) that α'/α will be between about 2 and 1.2. Thus the rigid former would not be able to confer much benefit. Sometimes copper is plated into a deep screw thread. This system does not appear to be wise, because the penetrating field will urge the current to concentrate at the bottom of the groove and thus alter the effective diameter of the coil by an amount which depends on the temperature. It is true that the contribution from the field inside the copper is small, since the plating is thin. But the arrangement approaches that of a strip winding on edge, considered and forbidden at the end of IV (C). Another construction is to copperplate the outside of a thin ceramic tube and subsequently cut a screw thread through the plating. This system seems a wise one, for the conductor is now a very thin strip on the surface of the tube. The internal inductance must be an extremely small fraction of the whole, and moreover will not decrease appreciably till the frequency is in the ultra-high spectrum. It would seem that this construction is capable of giving a coil whose temperature coefficient will differ insensibly from that of the former.

We have seen that $R_n \times L_{in}$ tends to be constant. Hence a resistance-frequency curve of a coil should give valuable information about its temperature coefficient. If at a given frequency the resistance has not increased by more than about 5 per cent, then it may be expected that the temperature coefficient will be sensibly that of linear expansion. If the resistance has increased considerably, then this increase should be at least sixteenfold (Z=45) in order that there may be a chance that α'/α is not greater than about 1.2.

FILAMENT DESIGN FOR HIGH-POWER TRANSMITTING VALVES*

By

J. J. VORMER

(Staatsbedrijf der P.T.T. Radiolaboratorium, 's-Gravenhage, Netherlands)

Summary—It is shown that the filament of a transmitting valve can be designed so that its life is a multiple of the accepted normal value. Not only do fewer service interruptions occur but the total operating expense decreases. The cost of power is computed at a reasonable wholesale rate.

IGH-POWER transmitting valves generally contain a directly heated tungsten cathode and two factors, surface area and temperature, determine the emission of such a filament.

For an existing type of valve, the length of the filament will be fixed and only the diameter and temperature may be varied. A change in these variables has only a slight influence upon the electrical properties of the valve. For a given total emission, the temperature and diameter vary inversely.

In this paper only the filament life and power will be considered. The power supplied to other electrodes will be independent of the filament for a given emission. It is assumed also that the manufacturing cost of a valve is practically independent of the diameter of the filament.

If the filament is thin, temperature must be high; the life will be short and the operating expense will be influenced largely by the cost of the valve whereas if the filament is thick, the more important factor will be the cost of power to heat it. An optimum value of diameter should give the lowest total operating expense.

The operating expense can be written as follows:

$$B = \frac{A}{L} + Wlc \tag{1}$$

in which

B = total hourly operating expense

A =price of the valve

L = life

W=power consumption in kilowatts per centimeter of filament length

l=filament length in centimeters

c=power cost per kilowatt-hour.

* Decimal classification: R331. Original manuscript received by the Institute, October 14, 1937.

For a given type of valve, A and l are constant while L and W are functions of the filament diameter d and temperature T.

We may substitute for T the factor I_s , which is the emission per centimeter of filament length and is a function of T and d. B may then be written as a function of d and I_s . This is a practical advantage as, for a specified type of valve, I_s may be computed directly from the total emission and effective length of the filament.

If the functions $L=f_1$ (I_s, d) and $W=f_2$ (I_s, d) are substituted in (1), the desired minimum in B can be found by means of the equation dB/dd=0. No difficulties are encountered in determining the relationship $W=f_2$ (I_s, d) . Data are available on the filament current and voltage and the emission current per centimeter as functions of T and d. These data from different authors are in satisfactory agreement. Fig. 1 is a family of curves plotting W as a function of d with I_s as a parameter.

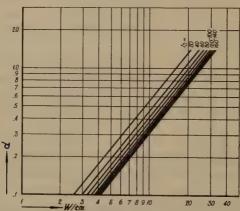


Fig. 1—Filament diameter plotted against heating power for various values of emission.

Determining the function $L=f_1$ (I_s,d) is less simple because life is not sharply defined. There is little agreement even as to the factors which influence it. Rukop¹ has published data in which life is plotted against the emission per watt of filament-heating power, with d as a parameter. The filament current was maintained constant and the end of life indicated by the burning out of the filament. The composition of the filament was not stated.

From experiments made by Simon² it may be concluded that L is proportional to d^2 . This is in fair agreement with the curves of Rukop

J. Zenneck and H. Rukop, "Lehrbuch der Drahtlosen Telegraphie," fifth edition, (1925), pp. 530-532.
 Simon, Handbuch der Experimentalphysik, vol. 13, II, (1928).

if very thin wires are not considered. Moreover Rukop mentions that under normal conditions the wire is ruined by recrystallization.

Contrary to this the "Marconi Company" states: "Experience has shown that a filament usually burns out when its evaporation has reduced its diameter by 10 per cent. As the rate of evaporation is independent of the diameter, it therefore follows that the life of a filament, when run at a definite temperature, is proportional to its diameter."

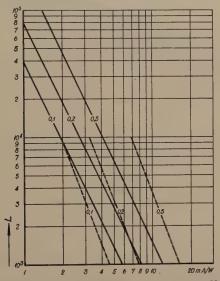


Fig. 2—Filament life plotted against emission for various filament diameters. The dashed graphs are for Rukop and the solid lines are the Marconi data.

The data of Rukop and Marconi are combined in Fig. 2 and illustrate that for high-power valves the differences in predicted life are considerable. For moderately thick wires Rukop's figures are a multiple of Marconi's. To check these figures in practical circumstances the life of a great number of transmitting and rectifying valves, in service at the transmitting stations of the Netherlands P.T.T., was determined. The valves were used normally in the transmitters and filament voltage was kept practically constant. This test showed that under these conditions small transmitting and rectifying valves have lives between the values predicted by Rukop and Marconi, while for high-power valves the attained lives were more in agreement with the data of Marconi. So the data of Marconi and Rukop may be considered as useful limits and calculations of the operating expense are made for both cases.

³ R. le Rossignol, and E. W. Hall, "The development of large radio transmitting valves," *Marconi Rev.*, no. 61, pp. 6-29; July-August; (1936).

Marconi and Rukop both show L as a function of T and d. From accepted data the emission per centimeter length I_s can be derived for different diameters and temperatures, and the function $L=f_1$ (d, I_s) determined. Figs. 3 (Marconi) and 4 (Rukop) show L=f (d) with I_s

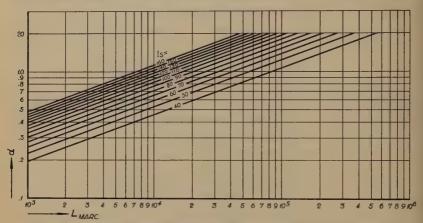


Fig. 3—Diameter plotted against life for various values of emission as reported by Marconi.

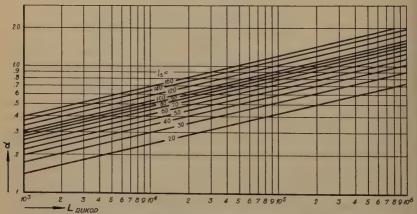


Fig. 4—Diameter plotted against life for various values of emission as reported by Rukop.

as a parameter. In these figures the data of both authors were extrapolated considerably whereby it was assumed that if d is constant, the relation between L and T is logarithmic, and that if d>0.5 millimeter, L is proportional to d^2 in the data of Rukop.

The graphs of Figs. 1, 3, and 4 are now presented mathematically. For this purpose it is assumed that they are composed of parallel straight lines which involves an error of not greater than a few percent.

The graphs of Fig. 1 become

$$W = 8.2617d^{0.78729}I_s^{0.21852} (2)$$

while the formulas for the life graphs are

$$L_{\text{Marc}} = 46.38 \cdot 10^6 \cdot d^{2.6584} I_{\bullet}^{-1.7289} \tag{3}$$

and

$$L_{\text{Rukop}} = 1500.4 \cdot 10^6 \cdot d^{4.1011} I_s^{-2.0300}$$
 (4)

By substituting (2) and (3) in (1) the operating expense per hour is found as a function of d and I_s ,

$$B_{\text{Marc}} = P_{\text{Marc}} \cdot d^{-2.6584} + Q d^{0.78729} \tag{5}$$

in which formula

$$P_{\text{Mare}} = \frac{A}{46.38} \cdot 10^6 I_s^{1.7289} \tag{6}$$

while

$$Q = 8.2617dl \cdot 10^{-3} I_{s}^{0.21852}. (7)$$

In the case of Rukop using (2), (4), and (1)

$$B_{\text{Rukop}} = P_{\text{Rukop}} d^{-4.1011} + Q d^{0.78729}$$
 (8)

in which formula

$$P_{\text{Rukop}} = \frac{A}{1500.4} \cdot 10^{-6} I_s^{2.0300}. \tag{9}$$

As the operating expense becomes a minimum if dB/dd=0 the most favorable wire diameter may be found. This optimum diameter, according to Marconi, is

$$d_{\text{opt Marc}} = \left[\frac{2.6584 P_{\text{Marc}}}{0.78729Q} \right]^{0.29021} = R_{\text{Marc}}^{0.29021}$$
 (10)

and according to Rukop

$$d_{\text{opt Rukop}} = \left[\frac{4.1011 P_{\text{Rukop}}}{0.78729 Q} \right]^{0.20047} = R_{\text{Rukop}}^{0.20047}.$$
 (11)

The corresponding optimum lives are

$$L_{\text{opt Maro}} = 46.38 \cdot 10^{6} I_{s}^{-1.7289} R_{\text{Maro}}^{0.77149}$$
 (12)

and

$$L_{\text{opt Rukop}} = 1500.4 \cdot 10^{6} I_{s}^{-2.0300} R_{\text{Rukop}}^{0.82215}. \tag{13}$$

OPTIMUM DIAMETER

These formulas are used to calculate the optimum filament diameter of two valves of existing types; i.e., a high-power air-cooled transmitting valve V_A and a water-cooled transmitting valve V_W . The cost of power is taken as f.0.03 per kilowatt-hour. The air-cooled valve V_A has the following characteristics:

filament current=16 amperes filament voltage=16 volts anode voltage=4000 volts anode dissipation=750 watts length of filament=24 centimeters total emission=1.5 amperes price=f.580.

If the cold ends of the filament are taken into consideration I_s =85 milliamperes per centimeter.

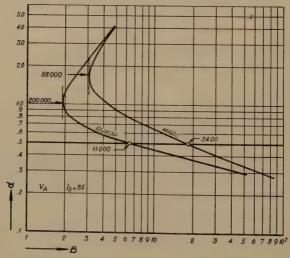


Fig. 5—Filament diameter plotted against operating expense for a sample radiation-cooled valve.

By means of (5) and (8) Fig. 5 is constructed. From (10) and (11) the optimum diameter is found.

 $d_{\text{opt Marc}} = 1.70 \text{ millimeters and } d_{\text{opt Rukop}} = 1.02 \text{ millimeters.}$

⁴ The florin is worth approximately 55 cents.

The corresponding optimum lives are found by (12) and (13).

 $L_{\rm opt\ Marc} \approx 88,000 \ {
m hours\ and} \ L_{\rm opt\ Rukop} \approx 200,000 \ {
m hours}.$

The valve V_A normally has a filament of 0.5 millimeter diameter. Its life will lie between 3400 and 11,000 hours according to (3) and (4) with the smaller value being most probable. So a filament of optimum diameter gives not only an enormous increase of life but decreases the operating expense by a factor of 4 to 6. The characteristics of the water-cooled valve V_W are:

filament current=77 amperes filament voltage=21.5 volts anode voltage=6000-12,000 volts anode dissipation=12,000 watts length of the filament=74 centimeters (2 wires in parallel) total emission=10 amperes price=f.1200.

Taking into account the cold ends of the filament I_s =160 milliamperes per centimeter. The relation between d and B is again found

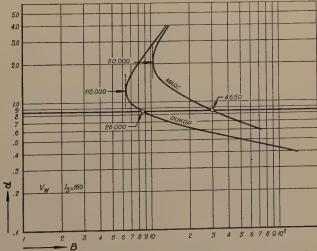


Fig. 6—Filament diameter plotted against operating expense for a sample water-cooled valve.

by means of (5) and (8) and is shown in Fig. 6. The optimum diameters as found from (10) and (11) are

 $d_{
m opt~Marc}=2.08$ millimeters and $d_{
m opt~Rukop}=1.23$ millimeters while the corresponding optimum lives are

 $L_{
m opt\ Mare} pprox 50,000 \ {
m hours\ and} \ L_{
m opt\ Rukop} pprox 115,000 \ {
m hours}_{
m o}$

The valve V_w is manufactured with a filament of 0.85-millimeter diameter and a life between 4650 and 26,000 hours.

The data of Rukop show that the operating expense for the standard filament amounts to f.0.085 while for the optimum diameter this figure is about f.0.064. Experience shows however that a life of not much more than 4650 hours will be obtained for the standard filament so the data of Rukop are not applicable for this range. It may, therefore, be expected that the operating expense can be reduced to about one third.

OPTIMUM POWER COST

The formulas permit also the determination of the cost of power for which a given diameter is optimum. Using (10) and (11) Q can be written

$$Q = \frac{2.6584 P_{\text{Marc}}}{0.78729} \cdot d_{\text{opt Marc}}^{-3.4458} \tag{14}$$

and

$$Q = \frac{4.1011 P_{\text{Rukop}}}{0.78729} \cdot d_{\text{opt Rukop}}^{-4.9883}. \tag{15}$$

In these formulas Q is the only factor containing the price of power c. If the diameters used by the manufacturer are optimum, it follows from (14) and (15) that the price per kilowatt-hour for V_A lies between f.1. and f.1.89 and for V_W between f.0.06 and f.0.30, the higher figures being the more reliable as they are derived from the Marconi data.

Practical Results

The calculations give only rough results as they are based on a number of approximations and extrapolations. Data concerning the life of a tungsten wire are fairly uncertain. Also, the corrections for the cold ends of the wire and the figures used for the number of watts per centimeter of filament length are only correct to a first approximation. The results of the calculation however appear to justify practical research.

The practical application should not present insurmountable difficulties. As the filament current is increased it will be necessary to choose suitable lead-in electrodes. The filament power being increased, some parts of the valve will attain a higher temperature. With aircooled valves this must be considered while with water-cooled valves this difficulty is probably less severe. If the valve V_A is provided with a filament of optimum diameter, i.e., 1.2 millimeters, the power consumption of the filament is multiplied by 2.6, an increase of about 400 watts. As the maximum anode dissipation is 750 watts, special precautions are necessary.

For the water-cooled valve V_W conditions are more favorable. The power consumption of the filament is multiplied by 2, an increase of about 1650 watts. The total anode dissipation being 12 kilowatts, this increase will probably cause no difficulties.

A number of transmitting valves were provided with thicker filaments. The diameter, originally 0.3 millimeter, was made 0.35 millimeter, this being the maximum permitted by the lead-in electrodes. The average life increased from 3492 to 7072 hours. With the data of Marconi the initial life is calculated at 3200 and the increased life at 4800 hours; with the data of Rukop these figures are 6000 and 14,000 hours. This increase in filament diameter has decreased the operating expense from f.0.052 to f.0.028 per hour.

Conclusion

It may be concluded that increasing the filament diameter by a comparatively small amount results in a considerable gain in life as well as in operating expense.

ACKNOWLEDGMENT

The author is indebted to Prof. Dr Ir N. Koomans, Chief of the Radio Laboratory, for his encouragement and stimulation.

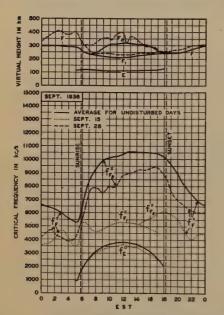
→>0≺**○**→·

CHARACTERISTICS OF THE IONOSPHERE AT WASHINGTON, D. C., SEPTEMBER, 1938*

Bı

T. R. GILLILAND, S. S. KIRBY, AND N. SMITH (National Bureau of Standards, Washington, D.C.)

ATA on the critical frequencies and virtual heights of the ionosphere layers during September, 1938, are given in Fig. 1. Fig. 2 gives the maximum frequencies which could be used for radio sky-wave communication by way of the F, F₂, and normal E layers. A



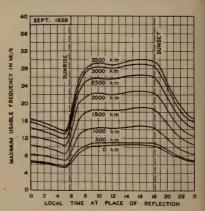


Fig. 2—Maximum usable frequencies for radio sky-wave transmission; averages for September, 1938, for undisturbed days, for dependable transmission by the regular F and F₂ layers.

Fig. 1—Virtual heights and critical frequencies of the ionosphere layers, September, 1938.

change toward winter conditions in the ionosphere was manifested by a decreased f_F , increase in the rate of rise of f_{F_2} in the morning, and a considerable increase in the f_{F_2} . Also, the daytime stratification of the

^{*} Decimal classification: R113.61. Original manuscript received by the Institute, October 10, 1938. This is one of a series of reports on the characteristics of the ionosphere at Washington, D.C. For earlier publications see Proc. I.R.E vol. 25, pp. 823-840; July, (1937), and a series of monthly reports beginning in Proc. I.R.E., vol. 25, pp. 1174-1191; September, (1937). Publication approved by the Director of the National Bureau of Standards of the U.S. Department of Commerce.

F layer was much less marked than in August, no well-defined f_{F_1} being observed, and the f_E was somewhat less than in August.

During September strong sporadic E reflections occurred seldom and not at very high frequencies; data are given in Table I.

TABLE I SPORADIC E Approximate Upper Limit of Frequency of the Stronger Sporadic E Reflections at Vertical Incidence

Hour, E.S.T.												
Date	00	01	02	03	04	05	06	07	08	09	10	11
Sept. 2 3 9			4.5	4.5	6	6	4.5	4.5	4.5		4.5 4.5	4.5

		Hour, E.S.T.											
Date	12	13	14	15	16	17	18	19	20	21	22	23	
Sept.					6	4.5	4.5						
8 28								$\frac{6}{4.5}$	4.5		8	6	

There were fewer ionosphere storms (Table II) and more sudden ionosphere disturbances observed in Washington (Table III) in September than in August.

TABLE II IONOSPHERE STORMS (APPROXIMATELY IN ORDER OF SEVERITY)

Date and	hr before	Minimum fr ^x before	Noon $f_{F_2}^x$	Magnetic	Ionosphere	
hour E.S.T.	sunrise (km)	sunrise (kc)	(ke)	00–12 G.M.T.	12-24 G.M.T.	character ²
14 (after 1600) 15 16 17 (to 0500) 27 (after 1800) 28 29 (to 0500) 26 (after 0400) 27 (to 1100)	376 366 302 390 356 320 360	3500 3650 4100 	<5400 8600 = 8800 8700	1.0 1.7 0.4 0.3 0.5 1.5 0.7 1.0	1.2 1.6 0.4 0.1 1.0 0.7 0.4 0.9 1.0	122 1 - 121-121 - 121-121-121
For comparison—Average for undisturbed days	295	5320	10380	0.2	0.2	0

Data on the degree of departure during September from the averages in Figs. 1 and 2 are given in Table IV. Information as to what time was disturbed is given in Table II.

¹ American magnetic character figure, based on observations of seven observatories.
² An estimate of the severity of the ionosphere storm at Washington on an arbitrary scale of 0 to 2, the character 2 representing the most severe disturbance.

	TABLE	III
SUDDEN	IONOSPHERE	DISTURBANCES

Dete	G.M	.T.	Locations of	Minimum			
Date 1938	Beginning of fade-out	Time of recovery	transmitters recorded	relative intensity ¹	Remarks		
Sept. 6	1730 1655	1920 1720	Ontario, Mass., D.C. Ontario, Mass., Ohio, D.C.	0.2 0.01			
20	1612	1700	Ontario, Mass., Ohio, D.C.	0.0			
20	1746	1900	Ontario, Mass.,	0.0	Terr. mag. pulse 1746–1835		
23 25	1654 1420	1755 1530	Ohio, D.C. Ontario, Ohio, D.C. Ohio, D.C.	$\substack{0.0\\0.05}$	1740-1855		

¹ Ratio of received field intensity during fade-out to average field intensity before and after, for station CFRX, 6070 kilocycles, 600 kilometers distant.

The days during which sporadic E-layer reflections were most prevalent at Washington are listed in Table I. The table shows the approximate upper limits of frequency at which strong sporadic E-layer reflections were observed at the hours listed. The observations were nearly continuous at 4.5, 6, 8, and 10 megacycles.

TABLE IV Percentage of Time for which the Ratio of f_F^x or $f_{F_2}^x$ to the Undisturbed Average Exceeded the Ratios at Heads of Columns

Ratio of critical frequency to average	0.4	0.5	0.6	0.7	0.8	0.9	1.0	1.1	1.2
Per cent of undisturbed time (582 hours)	100	100	100	100	100	90	55	7	0 0
Per cent of disturbed time (130 hours)	100	95	87	71	55	20	2	0	
Per cent of total time (712 hours)	100	99	97	94	91	77	45	6	

BOOK REVIEW

Les Communications Radio-Électriques (Radio Communications), by Henri de Bellescize. Gauthier-Villars, Libraire Du Bureau Des Longitudes, de l'École Polytechnique, 55, Quai des Grands-Augustins, 55, Paris. 1936. 220 pages. Price, 40 francs.

This book is devoted to the exposition of the principles underlying the engineering determination of dependable commercial communications in all of the recognized services, and of a more effective use of communication channels.

The first chapter covers those factors in receivers affecting the fidelity of the output signal to that arriving at the receiving antenna. The main factors of this nature discussed are the various forms of intermodulation, the magnitude and frequency of the local oscillator voltage, and the different methods of rectification.

In the second chapter (and appendix) are discussed the methods of evaluating the factors causing variations in the signal during its transmission through space, and as affected by disturbances, both natural and artificial. These methods are, of course, statistical in nature, and hence the author has prefaced his account with a brief development of those portions of the theory of probability useful for the computation of "distribution" curves, i.e., those in which an ordinate gives the probability that the corresponding abscissa is exceeded. He then goes in some detail into the computation of such curves from recordings of signal variations due to one or more causes, such as atmospherics, selective fading, etc. The chapter is rounded out with generalities as to the definition of "service" and the determination of the power required for different services (telegraph, telephone, broadcast).

These two chapters make up approximately one half of the book. The next three chapters are devoted to the experimental methods which have been developed within the past 20 years for reducing the variations in the received signal; chapter 3 discussing frequency stability and synchronization, chapter 4 antifading devices, and chapter 5 protection against disturbances. The final chapter (6) gives some applications of the principles developed in chapters 1 and 2, utilizing the methods described in chapters 3 to 5 to the determination of ranges, powers, etc., for particular services in both the short- and the long-wave domains.

From the above brief description of its contents, it will be seen that the book traverses a field in which, while possessing a rich periodical literature, there is a paucity of comprehensive treatments available. In fact, the reviewer knows of no other book written from quite the viewpoint of this one. That it is an important viewpoint for those interested in the maintenance of the security of communications and to allocation engineers goes without saying, and to such the reviewer can recommend this book unqualifiedly. It is written with the clarity which we have come to associate with the best French scientific expositions, is amply documented, and really deserves the hackneyed and often abused recommendation that it fills a long-felt want.

*L. P. WHEELER

^{*} Federal Communications Commission, Washington, D. C.

BOOKLETS, CATALOGS, AND PAMPHLETS RECEIVED

The following commercial publications of radio engineering interest have been received. Your request for a copy of any item may be addressed to the *Proceedings* for forwarding to the issuing company. Please mention your business affiliation.

Public Address • • • David Bogen Company, Inc. Catalog P938A, 20 pages, $7\frac{5}{5} \times 10\frac{3}{4}$ inches, printed. Description of public address amplifiers and systems.

QUARTZ CRYSTALS • • • Bliley Electric Company. Catalog G-10, 16 pages, $8\frac{1}{2} \times 11$ inches, printed. Quartz plates and holders for receiver and transmitter use.

Radio-Frequency Switches • • • Communication Products, Inc. 2 pages, $8\frac{1}{2} \times 11$ inches, lithograph. Describes 3 ceramic-insulated switches for low- and moderate-power service.

RADIO School • • • Capitol Radio Engineering Institute. Catalog, 48 pages +cover, 7½×10 inches, printed. Description of equipment and courses.

REMOTE ANTENNA AMMETER • • • Victor J. Andrew. Bulletin No. 88, 2 pages, $8\frac{1}{2} \times 11$ inches, printed. Specifications on a new indicating ammeter for use withhigh vertical broadcast radiators where the danger of breakdown due to lightning must be minimized.

RESISTORS • • • Clarostat Manufacturing Company. Printed. A series of bulleting giving engineering data on the Clarostat line of resistors.

Transformers • • • United Transformer Corporation. 16 pages, 8½×11 inches, printed. Catalog of audio and power transformers.

Tube Data (RCA) • • • RCA Manufacturing Company, Inc. Application Notes. No. 97, 3 pages, $8\frac{3}{8} \times 10\frac{7}{8}$ inches. "On a self-balancing phase-inverter circuit."

Tube Data (raytheon) • • • Raytheon Production Corporation. Engineering-Service release. 1 page. $8\frac{1}{2} \times 11$ inches. Tentative data on type 2A4G, an argon-filled filament-type thyratron.

VACUUM TUBE VOLTMETER • • • Ballantine Laboratories, Inc. Bulletin No. 2, 4 pages, $8\frac{1}{2} \times 11$ inches, printed. Description and operating characteristics of the Model 300 a-c voltmeter.

VIBRATORS • • • Electronic Laboratories, Inc. Folder AA, 8 pages, $8\frac{1}{2} \times 10\frac{3}{4}$ inches, lithograph. Descriptions of heavy-duty vibrators and polarity changers.

WIRE AND CABLES • • • Belden Manufacturing Company. Catalog No. 800, $8\frac{1}{2} \times 10\frac{1}{2}$ inches, 8 pages, lithograph. Lists the Belden line of wire and cables for communication-engineering applications.

WIRE STRIPPER • • • Ideal Commutator Dresser Company. Folder WSB, $8\frac{1}{2} \times 11$ inches, 6 pages, printed. Description of various types of equipment for stripping the insulation from covered wires and cables.

CONTRIBUTORS TO THIS ISSUE

Bingley, Frank James: Received B. Sc. degree (Special) Hons. Mathematics, 1926; B. Sc. degree (Special) Hons. Physics, 1927, London University. Technical assistant, Baird Television Company, London, 1927–1928; chief engineer, Baird Television Company, New York, 1928–1930; television research, Philoo Radio and Television Corporation, 1931 to date. Associate member, Institute of Radio Engineers, 1934; Member, 1936.

Chase, Rush F.: Born January 20, 1914, at Hot Springs, South Dakota. Received B.S. degree in electrical engineering, University of Washington, 1937. Radio serviceman, General Electric Supply Corporation, summer, 1936; engineer, Communication Engineers, Inc., summer, 1937; builder of office interphones, Radio Laboratories, Inc., fall, 1937; operator of Madison Park Electric Company, January, 1938, to date. Student member, Institute of Radio Engineers, 1937.

Freeman, Robert L.: Born February 1, 1909, at Cleveland, Ohio. Received A.B. degree, 1931; E. E. degree, 1933; Ph.D. degree, 1934, Stanford University. Receiver development, Crosley Radio Corporation, 1934–1937; transmitter engineer, Farnsworth Television, Inc., 1937; engineer, Hazeltine Service Corporation, 1937 to date. Member, Sigma Xi. Student member, Institute of Radio Engineers, 1932; Associate, 1933.

George, Ralph W.: Born July 19, 1905, at Jamestown, North Dakota. Received B.S. degree in electrical engineering, Kansas State College, 1928. Field laboratories, R.C.A. Communications, Inc., 1928 to date. Associate member, Institute of Radio Engineers, 1931.

Gilliland, T. R.: See Proceedings for January, 1938.

Ginzton, Edward Leonard: Born December 27, 1915, in Russia. Received B.S. degree, 1936; M.S. degree, 1937, University of California; E.E. degree, 1938, Stanford University. Assistant in teaching and research, Stanford University, 1937 to date. Member of Tau Beta Pi, Sigma Xi, Eta Kappa Nu, and Student member, A.I.E.E. Nonmember, Institute of Radio Engineers.

Ito, Yoji: Born March 5, 1901, at Tiba, Japan. Received B.S. degree in electrical engineering, Tokyo Imperial University, 1924; Dr.-Ing. degree in electrical engineering, Dresden Technical High School, 1929; M.S. degree in electrical engineering, Tokyo Imperial University, 1936. Engineering officer, Imperial Japanese Navy, 1924 to date. Now ordnance commander, Imperial Japanese Navy. Nonmember, Institute of Radio Engineers.

Kirby, S. S.: See Proceedings for January, 1938.

Kroger, Fred H.: Engineer, R.C.A. Communications, Inc., 1920 to date. Member, Institute of Radio Engineers, 1913; Fellow, 1928.

Moullin, E. B.: Born 1893. Graduated in engineering, Cambridge University, England, 1916. University lecturer, Cambridge University, 1920–1929; reader in

engineering science, Oxford University, 1929 to date. Member, Radio Research Board, Department of Scientific and Industrial Research of Great Britain, and of various subcommittees of the same Board. Author of "Radio Frequency Measurements," "Principles of Electromagnetism," and "Spontaneous Fluctuations of Voltage." Member, Institution of Electrical Engineers, London. Member, Institute of Radio Engineers, 1927.

Reymer, S. E.: See Proceedings for January, 1938.

Sinkawa, Hirosi: Born November 13, 1909, at Tokyo, Japan. Received B.S. degree in electrical engineering, Waseda University, 1933. Engineer in Naval Technical Research Department, 1933 to date. Nonmember, Institute of Radio Engineers.

Smith, J. Ernest: Born July 9, 1908, at Carrollton, Missouri. Received B.A. degree in mathematical physics, Jamestown College, 1932; graduate scholar, California Institute of Technology, 1933. Instructor in communication engineering, RCA Institutes, 1934–1938. Engineering department, central office laboratory, R.C.A. Communications, Inc., 1935 to date. Associate member, Institute of Radio Engineers, 1937.

Smith, Newbern: See Proceedings for January, 1938.

Tani, Keikitiro: Born April 24, 1895, at Kanazawa, Japan. Received B.S. degree in electrical engineering, 1919; M.S. degree in electrical engineering, Tokyo Imperial University, 1932. Engineering officer, Imperial Japanese Navy, 1919 to date. Now ordnance Captain, Imperial Japanese Navy, engineer in Patent Office, and secretary of the Radio Research Committee of National Research Council of Japan. Nonmember, Institute of Radio Engineers.

Vormer, J. J.: Born October 24, 1901, at The Hague, Holland. Received degree of Electrotechnical Engineer, Technische Hoogeschool, Delft, 1925. Engineer, Radio Laboratory, Netherlands States Telegraph, 1925 to date. Nonmember, Institute of Radio Engineers.
